PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : G01N 33/53, 33/566, 33/543, C12Q 1/00, A61K 38/00 — A1

(11) International Publication Number: WO 00/63694

(43) International Publication Date: 26 October 2000 (26.10.00)

(21) International Application Number: PCT/US00/09714

(22) International Filing Date: 12 April 2000 (12.04.00)

(30) Priority Data:

60/129,417	15 April 1999 (15.04.99)	US
60/132,595	5 May 1999 (05.05.99)	US

(71) Applicant *(for all designated States except US)*: UNIVERSITY OF VIRGINIA PATENT FOUNDATION [US/US]; 1224 West Main Street, Suite 1–110, Charlottesville, VA 22903 (US).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: HAYSTEAD, Timothy, A., J. [GB/US]; 1044 Hayrake Lane, Charlottesville, VA 22903 (US).

(74) Agent: BREEN, John, P.; University of Virginia Patent Foundation, Suite 1–110, 1224 West Main Street, Charlottesville, VA 22903 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

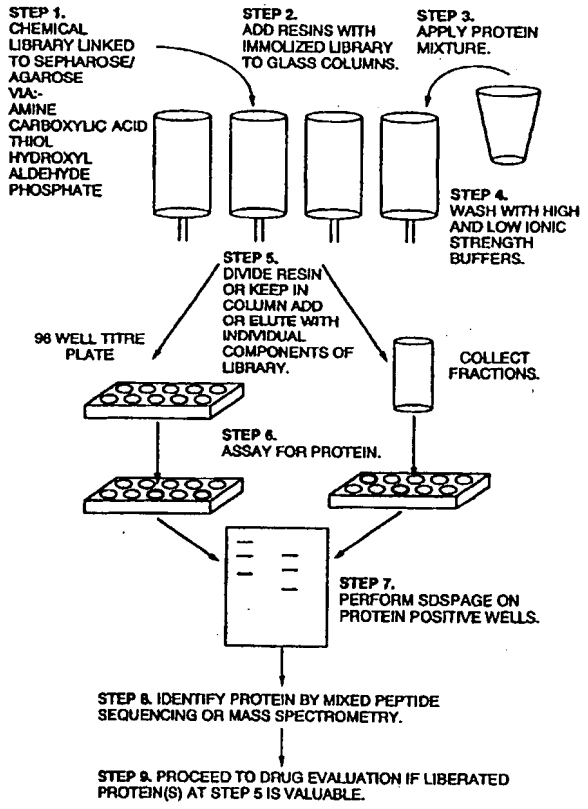
Published

With international search report.

(54) Title: PROTEOME MINING

(57) Abstract

The present invention relates to a method and apparatus for screening diverse arrays of materials for bioactive compounds. In particular, techniques for rapidly characterizing compounds in arrays of materials in order to discover and/or optimize new materials with specific desired properties are provided. The figure represents one of the embodiments of the current invention method for isolating bioactive compounds from a complex mixture of proteins using an immobilized combinatorial library.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

Proteome Mining

US Government Rights

This invention was made with United States Government support under Grant No. DK52378A, awarded by the National Institutes of Health. The United States Government has certain rights in the invention.

Field of the Invention

The present invention is directed to high throughput screening of proteomics for identifying bioactive compounds. More particularly the invention is directed to the isolation of novel herbicides, antibiotics, antifungals, antivirals, insecticide or pharmaceuticals based on their interactions with target molecules.

Background of the Invention

The animal, plant, prokaryotic and viral kingdoms contain within them a vast array of genes that express 100,000's of distinct proteins whose biological function is essential life. The number of genes contained with in a particular organism varies greatly. Generally, the simpler the organism the fewer the total number genes. For example, completion of the yeast genome shows that these organisms have about 8300 genes, the complete *C. elegans* genome contains about 18,000 genes, and the human genome is estimated to contain about 100,000 distinct genes. Each gene encodes a specific protein which has a predetermined essential function for the over all survival of the organism. Collectively, given the biodiversity that exists on earth, the numbers of distinct genes that exist in nature is likely to number in the billions.

Obviously not all of the proteins expressed by the genes of an organism are likely to be of importance to man. Indeed the number of genes that are likely to express proteins of commercial or medical value is a tiny fraction of this vast biodiverse gene pool. Methods therefore that allow one to rapidly and effectively screen large numbers of proteins within this pool for valuable proteins are of great importance. In the case of the human genome it has been estimated that approximately 4000 of its genes are responsible for the causes of non-pathogen

induced human disease. In short this means that human tissue contain 4000 proteins of potential medical and commercial value. The same analogies can be made to other species. For example if one was screening for a new antibacterial agent, one would be looking for bacterial protein targets that were peculiar to the particular bacterial strain of interest. In the case of bacteria this has traditional been enzymes involved in the synthesis of the bacterial cell wall. The classic example of a drug that is selective for bacteria is penicillin which inhibits an essential enzyme required for synthesis of the bacterial cell wall. Humans do not possess any of the enzymes that make bacterial cell walls. One cannot simply target any bacterial protein when searching for new antibiotics, this is because even though there are many differences between humans and bacteria, a significant portion of the bacterial genome encodes proteins of similar structure or function as found in humans. Drugs that inhibited proteins with a common function in both organisms are unlikely to discriminate between the two species.

To identify new drugs or commercially important bio-active molecules one needs methods that have the ability to encompass entire species genomes and immediately identify candidate proteins of importance. The present invention is directed to a method of identifying compounds that selectively interact with important biological components. This selective interaction is an essential element that makes a particular chemical have medical or commercial value. Without selectivity a compound has no bio-active value; selectivity is the single most important factor in all drug, antibiotic, antifungal, antiviral, insecticide and herbicide action.

The selectivity of a valuable bio-active compound in 99% of all cases is based on its interactions with one or more specific proteins contained within the target cell or organism of interest. One or two percent of valuable bio-active molecules maybe directed towards non-protein targets such as DNA, RNA, lipids or sugars. Without exception a valuable bio-active chemical interacts with its protein target in a highly specific manner. The target protein will contain on its surface a domain or pocket that binds the chemical with high affinity. This domain or pocket is unique to the target and not the several 1000 proteins that may also be contained with in the cell expressing the target protein.

In most cases the binding site for the chemical on the target protein is

important to the biological activity of the protein. The binding site may be required for enzymatic activity, or be site of hormone interaction (e.g. a receptor) or a binding site for an all osteric regulator. When a chemical binds to one of these sites and affects the biological activity of the target protein it invariably contains structural components that resemble the natural biological ligand that normally occupies the site. The affinity of the interaction of the chemical for the target protein generally increases as the overall structural components of the chemical mimic the natural ligand. In some cases bio-active molecules fit better into the natural ligand-binding site than the natural ligand itself. These types of molecules are likely to have extremely potent bio-activity. If these molecules possess structural features that prevent their metabolism then this increases their bio-activity even greater.

One of the primary mechanisms for identifying bio-active chemicals of medical or commercial value is to screen large combinatorial libraries with some form of an enzymatic or biological assay. Combinatorial libraries can be extremely diverse and contain many hundred thousands of distinct molecules of known or unknown structure. They can be derived from very diverse sources, including plant extracts, animal extracts, soil samples, bacteria, fungi, chemical industry byproducts etc. Theoretically, these libraries contain within them molecules of every conceivable shape and form. However, like the proteins to be targeted, only a small percentage of these libraries contain molecules that have important bio-activity.

Prior high through put screens for drug discovery begin with a disease, the choice of which is invariably determined by potential market size. The etiology of the disease is first defined by basic research to determine likely underlying cause. This research identifies potential protein targets that may be useful drug targets; e.g. receptors or enzymes. The purified receptors or enzymes are then used to screen chemical libraries for agents either bind, inhibit or activate. Similar approaches are also used to screen for anticancer drugs. Transformed cell lines are used to screen large chemical libraries that may contain compounds that revert them to their normal phenotype or kill the cancer. Further investigation is then used to identify the molecular mechanisms by which active compounds from these screens bring about their cellular effect.

Therefore in the traditional search for a new bio-active compounds one

begins with a specific biological problem in mind. For example a particular pharmaceutical house maybe focused on discovering new antihypertensive agents. Its decision to enter into the screening process is always based on the disease market size. Thus the search for drugs that would treat small populations of afflicted individuals is unlikely to happen in the private sector. In the case of a new antihypertensive agent for example, a drug screen will generally begin with an assay that includes a specific receptor or enzyme that has important functions in the regulation of blood pressure. One then has to hope that the libraries one screens contain bio-active molecules that selectively effect the proteins selected in the assay. Once candidate chemicals are identified one then has to demonstrate that these compound act selectively and predictably for the targeted protein in the assay and not others. Thus in the initial stages one ends up with many false positives which must be eliminated in a second round of screening because the entire expressed genome was not taken into account in the first instance. The invention described herein eliminates this problem at the start because it encompasses both the diversity contained within the chemical library to be screened, with the diversity of the expressed genome itself in one step.

The selectivity is achieved in the analysis following sequencing of the targeted proteins. A decision as to whether a particular protein/chemical interaction is likely to have commercial and medical value is made during the last stages of analysis. Therefore, in addition to identifying bio-active agents that have commercial value the screen of the present invention does not exclude compounds that may have humanitarian value. This is because we could conceivably identify agents that bind to proteins important in the pathology of obscure diseases with small patient populations.

Finally, the present invention can readily cross platforms with no change in protocol or equipment. There is no difference in screening procedures for herbicides, antibiotics, antifungals, antivirals, insecticide or pharmaceuticals. All one changes is the expressed genome (proteome) that is to be screened. One can even use the same libraries for each screen; i.e. a library that did not yield any useful pharmaceutical agents may contain a useful herbicide.

Summary of the Invention

The present invention is directed to compositions and methods for identifying bioactive compounds. Advantageously, the present method of identifying bioactive compounds utilizes both the diversity contained within a chemical library to be screened, with the diversity of the expressed genome itself in one step to maximize the efficacy of the screening procedure. The method comprises the steps of contacting an immobilized combinatorial library with the protein members of a proteome, characterizing the proteins that interact with members of the library to identify those proteins having important biological value, and isolating the corresponding compound from the library that interacts with a protein having important biological value.

Brief Description of the Drawings

Fig. 1 is a diagramatic representation of the steps used in accordance with one embodiment to isolate bioactive compounds from a complex mixture of proteins (proteome) using an immobilized combinatorial library.

Fig. 2 is a diagramatic representation of the steps used in accordance with one embodiment to isolate bioactive compounds from a complex mixture of proteins (proteome) using an immobilized combinatorial library.

Fig. 3 is a diagramatic representation of the steps used in accordance with one embodiment to isolate bioactive compounds from a complex mixture of proteins (proteome) using an immobilized combinatorial library.

Fig. 4 is a diagramatic representation of the steps used in accordance with one embodiment to identify cell surface receptors and their peptide ligands en masse from a predetermined cell type.

Fig. 5 represents the stained SDSPAGE results from characterization of proteins isolated from rabbit skeletal muscle through the use of gammaphosphate linked ATP-Sepharose. Rabbit skeletal muscle extract was prepared from 350 g of tissue (w/w) and passed over 50 mls of gamma phosphate linked ATP-Sepharose containing 10 umols/ml of linked ATP. Following washing, the column was eluted sequentially with NADH, AMP, ADP and ATP and fractions collected (10mls). Column fractions were separated by SDSPAGE then transfered to PVM and stained with amido black. Proteins 1-17 were identified by mixed peptide sequencing (see

Table 1).

Fig. 6 represents the stained SDSPAGE results from geldanamycin released muscle extract proteins. ATP-Sepharose was loaded with skeletal muscle extract and eluted successively with the indicated concentrations of geldanamycin, followed by 10 mM ATP. Peak fractions (20ul of 1.0) were analyzed by SDSPAGE and silver staining or transferred to PVM for identification by peptide sequencing. Numbers indicate that proteins that wee identified on the PVM membrane: 1. HSP90 and proteolytic fragments of HSP90; 2. purine synthetase (ADE2); 3. myosin light chain kinase; 4. phosphorylase kinase; 5. p98 glucose indued kinase; 6. HSP70; arginine succinate synthetase; 7. glutamate dehydrogenase; 8. glutamate ammonium ligase; 9. glutathione sythetase; 10. aldehyde dehydrogenase; 11. MAPK; 12. GAPDH; 13. PKA

Detailed Description of the Invention

In describing and claiming the invention, the following terminology will be used in accordance with the definitions set forth below.

As used herein, "nucleic acid," "DNA," and similar terms also include nucleic acid analogs, i.e. analogs having other than a phosphodiester backbone. For example, the so-called "peptide nucleic acids," which are known in the art and have peptide bonds instead of phosphodiester bonds in the backbone, are considered within the scope of the present invention.

As used herein, bioactive compounds include any compound that is capable of inducing an effect on a living cell or organism. Bioactive compounds include but are not limited to pharmaceuticals, hormones, chemotherapeutics, nucleic acids and the like.

As used herein the term "proteome" relates to a complex mixture of proteins that are derived from a common source, such as an extract isolated from a particular cell or tissue. For example a human proteome represents a mixture of proteins isolated from human cells. The category can be further defined by specifying a particular cell/tissue source for the proteome (i.e. a human myocardial tissue proteome represents all the proteins isolated from human myocardial tissue).

As used herein the term "combinatorial library" relates to a collection

of compounds. The combinatorial library can be a biological synthesized library that comprises nucleic acid sequences that include a common vector sequence (allowing for replication of the library in a host species) and a protein encoding region. The biological synthesized library can be further provided with regulatory elements that allow for the expression of the encoded proteins (i.e. an expression library). Chemical libraries are collections of compounds that were isolated from a natural source or were synthesized in a laboratory using chemical or biological processes. A "combinatorial chemical library" is a collection of compounds created by a combinatorial chemical process, wherein the compounds of the combinatorial chemical library have a common scaffold with one or more variable substituents.

As used herein the term "solid support" relates to a solvent insoluble substrate that is capable of forming linkages (preferably covalent bonds) with soluble molecules. The support can be either biological in nature, such as, without limitation, a cell or bacteriophage particle, or synthetic, such as, without limitation, an acrylamide derivative, agarose, cellulose, nylon, silica, or magnetized particles.

As used herein the term "naturally-occurring" relates compounds normally found in nature. Although a chemical entity may be naturally occurring in general, it need not be made or derived from natural sources in any specific instance.

As used herein the term "non naturally-occurring" relates to compounds rarely or never found in nature and/or made using organic synthetic methods.

As used herein the term "functional analog" of a library compound/ligand relates to a compound that has a binding affinity for the same ligand as one of the members of the library, such that the functional analog will compete with the library component for binding to that ligand.

The present invention is directed to a novel method for the rapid identification of bioactive compounds, including but not limited to novel drugs, antibiotics, antifungals, antiviral, insecticide or herbicides. The overall strategy behind the invention is to screen complex protein mixtures with an immobilized library of compounds for proteins that bind specifically to components in the library. The bound proteins are then identified by protein microsequencing to determine if the

identified protein is therapeutically relevant. The protein has therapeutic relevance if the protein is known to be central to the development of a disease, is a metabolic enzyme unique to a particular microorganism, yeast, virus, or fungi, or is an enzyme peculiar to a type of insect, or an enzyme required for photosynthesis in a particular weed.

The advantage of the present screening methodology derives from the initial assumption that the entire genome is a potential drug target. The only decision prior to screening is to decide what proteome should be utilized; i.e. for drugs important in human disease, human tissue is the choice, for herbicide a plant species, for antibiotic, a bacterial strain.

In one aspect of the invention, systems and methods are provided for rapidly screening a combinatorial library for bioactive agents. The method is based on the identification of those library components that interact with proteins of a preselected proteome, wherein the proteome protein is a potential target for therapeutics. The method of identifying bioactive compounds comprises the steps of contacting a combinatorial library with the protein members of a proteome under conditions that allow for specific interactions between proteins of the proteome and the bound library. Proteins that interact with the immobilized library components are then isolated and analyzed to determine if the protein is interesting from a therapeutic standpoint. Those proteins that have therapeutic relevance are then used to identify the component of the immobilized library that interacts with the protein.

In one embodiment, the method of identifying bioactive compounds present in a compound library comprises the steps of contacting an immobilized compound library with the protein members of a preselected proteome, and washing the immobilized compound library with a buffered solution. In one embodiment the immobilized compound library comprises a column of particulate solid support, such as sepharose or agarose beads, that has the individual components of the compound library bound to the support, and the wash comprises the use of a low ionic or high ionic buffer. In one embodiment the column is washed with both a high ionic buffer and a low ionic buffer. After the solid support has been washed with buffer, the remaining bound proteins are released from the solid support by contacting the bound proteins with one or more individual members of the compound library or with

functional analogs of the library components. Alternatively, the bound proteins can be released from the immobilized library through the use of a chaotropic agent, including but not limited to detergents such as SDS, TritonX, sarkosyl, denaturants such as urea, or chelators such as EGTA or EDTA. The released proteins are then identified by protein sequencing or mass spectrometry; and the identity of the specific compounds of the compound library that bind to the released proteins is determined.

The libraries

Combinatorial libraries can be constructed using techniques known to the skilled practitioner to provides researchers vast number of chemical candidates to screen for potential bioactivity. In accordance with the present invention the library comprises a collection of compounds that are capable of specific binding to their target. For example, suitable library components include, but are not limited to peptides, proteins, carbohydrates, lipids, glycoproteins or nucleic acids.

Biologically synthesized combinatorial libraries have been constructed using techniques of molecular biology. These library components are expressed using bacteria or bacteriophage particles. For example, U.S. Pat. No. 5,270,170 and 5,338,665 to Schatz describe the construction of a recombinant plasmid encoding a fusion protein created through the use of random oligonucleotides inserted into a cloning site of the plasmid. In other biological systems, for example as described in Goedell et al., U.S. Pat. No. 5,223,408, nucleic acid vectors are used wherein a random oligonucleotide is fused to a portion of a gene encoding the transmembrane portion of an integral protein. Upon expression of the fusion protein it is embedded in the outer cell membrane with the random polypeptide portion of the protein facing outward. Thus, in this sort of combinatorial library the compound to be tested is linked to a solid support, i.e., the cell itself and the cell itself adheres to the cell culture substrate. The Goedell patent is incorporated herein by reference.

Similarly, bacteriophage display libraries have been constructed through cloning random oligonucleotides within a portion of a gene encoding one or more of the phage coat proteins. Upon assembly of the phage particles, the random polypeptides also face outward for screening. Such phage expression libraries are described in, for example, Sawyer et al., 4 Protein Engineering 947-53 (1991);

Akamatsu et al., 151 J. Immunol. 4651-59 (1993), and Dower et al., U.S. Pat. No. 5,427,908. These patents and publications are incorporated herein by reference.

While synthesis of combinatorial libraries in living cells has distinct advantages, including the linkage of the compound to be tested with its nucleic acid, there are clear disadvantages to using such systems as well. The diversity of a combinatorial library is limited by the number and nature of the building blocks used to construct it; thus modified or R-amino acids or atypical nucleotides may not be able to be used by living cells (or by bacteriophage or virus particles) to synthesize novel peptides and oligonucleotides. There is also a limiting selective process at play in such systems, since compounds having lethal or deleterious activities on the host cell or on bacteriophage infectivity or assembly processes will not be present or may be negatively selected for in the library.

Another approach to creating molecularly diverse combinatorial libraries employs chemical synthetic methods to make use of atypical or non-biological building blocks in the assembly of the compounds to be tested. Thus, Zuckermann et al., 37 J. Med. Chem. 2678-85 (1994), describe the construction of a library using a variety of N-(substituted) glycines for the synthesis of peptide-like compounds termed "peptoids". The substitutions were chosen to provide a series of aromatic substitutions, a series of hydroxylated side substitutions, and a diverse set of substitutions including branched, amino, and heterocyclic structures. This publication is incorporated by reference herein.

Alternatively, chemical synthetic methodologies can be used to create large diverse libraries of potentially useful compounds and permits the synthesis of compounds joined to a solid support of some kind or joined to an identifiable marker such as a flourescent tag. In accordance with one embodiment, the combinatorial library is chemically synthesized on solid supports in a methodical and predetermined fashion, so that the placement of each library member gives information concerning the synthetic structure of that compound. Examples of such methods are described, for example, in Geysen, U.S. Pat. No. 4,833,092, in which compounds are synthesized on functionalized polyethylene pins designed to fit a 96 well microtiter dish so that the position of the pin gives the researcher information as to the compound's structure. Similarly Hudson et al., PCT Publication No. W094/05394, describe methods for the

construction of combinatorial libraries of biopolymers, such as polypeptides, oligonucleotides and oligosaccharides, on a spatially addressable solid phase plate coated with a functionalized polymer film. In this system the compounds are synthesized and screened directly on the plate. Knowledge of the position of a given compound on the plate yields information concerning the nature and order of building blocks comprising the compound.

Another approach has been the use of large numbers of very small derivatized beads, which are divided into as many equal portions as there are different building blocks. In the first step of the synthesis, each of these portions is reacted with a different building block. The beads are then thoroughly mixed and again divided into the same number of equal portions. In the second step of the synthesis each portion, now theoretically containing equal amounts of each building block linked to a bead, is reacted with a different building block. The beads are again mixed and separated, and the process is repeated as desired to yield a large number of different compounds, with each bead containing only one type of compound. This methodology, termed the "one-bead one-compound" method, yields a mixture of beads with each bead potentially bearing a different compound. The compounds displayed in the surface of each bead can be tested for the ability to bind with a specific compound (i.e. a protein member of a proteome).

The libraries used in the present invention can be well defined, containing known mixtures of molecules, or the library can be one in which the chemical content is poorly defined.

In accordance with one embodiment the libraries are immobilized on a solid support. Biological material, including but not limited to proteins, carbohydrates, nucleic acids, lipids, glycoproteins can be bound to a solid surface using standard techniques known to those skilled in the art. In preferred embodiments the library compounds are linked through covalent bonds. The solid surface can be selected from any surface that has been used to immobilize biological compounds and includes but is not limited to polystyrene, agarose, silica or nitrocellulose. In one embodiment the solid surface comprises functionalized silica or agarose beads.

In accordance with one embodiment the components of a sample are bound to silica or agarose beads in separate reactions using different reaction reagents

and conditions to ensure that a diverse array of compounds are bound to the solid surface. Fractions of these separate reactions can then be combined to form a single affinity chromatography column. For example, a portion of the library can be reacted with an inert solid support (e.g. agarose, Sepharose, polystyrene or other chromatography beads) using standard protocols for linking primary amines to NHS, cyanogen bromide activated or maleimide activated resins. Many of these resins are commercially available, for example, Pharmacia CH-activated Sepharose. Another portion of the library is then reacted to an inert support that would select any molecules containing carboxylic acid residues. A commercially available resin in this case would be Pharmacia EAH-activated Sepharose. A third and fourth portion of the library could be linked through thiol (SH2), phosphate or aldehyde (CHO) containing residues. The goal is to link as many components as possible with in the library and in as many orientations as possible. The orientation of molecule when it is tethered is critical to its ability to interact with potential target proteins. Thus for some molecules reaction through primary amines may cause binding of a target protein to be sterically hindered because of the thether. However, tethering of the same molecule at a carboxyl residue may not hinder interaction with a target protein.

The entire library can be linked to the resin or portions of the library linked can be linked separately. One should aim to achieve as reasonably high a ligand concentration as possible per immobilized molecule. Ideally this should be between 10 nmol to 1 μmol/ml. In the case of libraries in which the chemical content is poorly defined two mixtures from the library are prepared, a mixture that is soluble in organic phase and a mixture that is soluble in aqueous phase. The same linking strategy is then employed for the preparation of the resins from the organic and aqueous soluble library members. Each resin is then placed into a chromatography column and equilibrated into the protein extraction buffer.

In another embodiment (shown in Fig 2), individual components in a library, or mixtures containing between 1-10 chemicals, are attached to beads separately in microtitre plate wells. Several 100 beads can be reacted in each well at the same time. One bead is then selected from each well and placed in a chromatography column. Again each ligand should be attached in multiple orientations. In cases where a target compound is identified that interacts with a large

number of proteins (e.g. ATP, see Fig 3) only the target compound is immobilized on the solid support. The bound proteins can be released by the addition of an excess of free target compound, a derivative of the target compound or a functional analog of the target compound (see Example 2).

The proteomes

The immobilized libraries are contacted with a proteome under conditions conducive to the formation of specific interactions between components of the proteome and components of the immobilized library. Choice of the proteome used depends on the problem being addressed. If one is looking to isolate a new drug for the treatment of a human disease the human proteome (i.e. human tissue) is the best choice. If one wishes to discover a new antibiotic then the target pathogen (e.g. a gram negative bacterium) of interest is the obvious choice. For an insecticide, the targeted insect and so on.

In accordance with one embodiment the proteome comprises a natural products library which represents a collection of natural products that have been recovered from biological material and have been determined to have biological activity. For example the natural products library may include a mixture of natural products wherein the mixture is known to induce a phenotypic change in a population of cultured cells.

The amount of starting tissue used to isolated the proteome proteins is critical and should be based on theoretical recovery of target proteins. For example, if one is interested in identifying drugs that may bind to signal transduction molecules one should take into account the amount of these proteins (or copy number) that may be in the cell. Many of these types of molecules may be expressed as low as 200 copies per cell. A quick calculation predicts that if one wanted to recover as much as 10 pmol of a particular protein that was expressed at 200 copies per cell one would require 12 grams of wet weight tissue. For high copy number proteins, such as metabolic enzymes, less tissue mass would be required to achieve 10 pmol of protein. One should also factor in potential losses due to inefficiency of extraction or proteolysis. Thus although 12 g of tissue may contain a total of 10 pmol of a target protein of interest one may only recover a small percentage of this in the initial screen.

Fortunately, modern protein detection and sequencing methods enable one to identify proteins in the femto molar range. However, increasing the starting tissue mass as much as possible will improve the odds of recovering sufficient protein for later identification.

Tissue from the chosen target proteome is ground and homogenized in buffers appropriate for solubilizing proteins and retaining their native conformations. This involves standard procedures utilized in most biochemical laboratories. Following clarification by centrifugation the portions of the extraction mixture are passed over the immobilized chemical library resins.

When interactions are sought between library compounds and cell surface proteins, it may be advantageous to investigate such interactions with the surface proteins in their natural state (i.e. embedded in the cell membrane). In accordance with one embodiment, the "proteome" represents the set of surface molecules displayed by cells cultured on a cell culture substrate, and may include proteins, glycoproteins, carbohydrates and lipids.

In accordance with one embodiment certain components of the proteome can be removed prior to contacting the immobilized compound library with the proteome. Components can be removed, for example, by fractionating the components based on molecular weight, electric charge and/or hydrophobicity. Alternatively, specific components can also be removed by ligand or antibody binding. Such methods allow the removal of protein components that do not warrant further investigation but are know to bind to certain components of the target compound library. In addition such prescreening allows for the removal or reduction of proteins that are expressed at high levels in the tissue used to generate the proteome.

The compounds of the proteome are placed in contact with the library component under conditions favorable for specific interactions between members of the two groups. The interaction may result in the alteration of a physical characteristic such as fluorescence, absorption, enzymatic activity, but typically the specific interaction involves binding of the two components to one another. In accordance with one embodiment the library components will be immobilized on a solid support and the proteome components will be solubilized or suspended in a solvent. The

solvent will then be incubated with the immobilied library for a time sufficient for specific interaction between the library components and the proteome components.

In accordance with one embodiment the library components are covalently linked to a particulate solid support and the particulate is used to form a column. In this embodiment the proteome solution/suspension is passed through the column to provide contact between the library components and the proteome components. Some of the proteome components will bind to the immobilized library components. The immobilized library is then washed with a buffered solution to remove any non-specifically bound proteins. In accordance with one embodiment the step of washing the library comprises the steps of washing with a high ionic strength buffer and a low ionic strength buffer to remove proteins that may be associated because of non-specific ionic or hydrophobic interactions..

Isolation and identification of target proteins with a defined immobilized library

In accordance with one embodiment the library comprises a defined set of compounds that have been covalently linked to sepharose/agarose beads (resin) via amine, carboxylic acid, thiol, hydroxyl, aldehyde or phosphate linkages and the beads are combined to form a column (see Fig 1). The protein mixture (proteome) is then applied to the column followed by washes of high and low strength buffers. In accordance with one embodiment the solutions/suspensions are allowed to percolate through the column based on gravity. Alternatively, additional force can be applied to speed the flow of the eluate through the column; for example the column can be spun in a centrifuge to enhance flow through the column.

After the column has been washed with the buffered solutions, the beads (resins) are either kept in the columns or removed and placed in equal amounts into 96 well microtitre plates (see step 5, Fig. 1). Maintaining the resins in a column has the advantage of potentially recovering more protein per library ligand, however, it has the disadvantage of being much slower procedure overall.

When using the microtitre plate approach, the number of plates and wells used depends upon the number of components that are in the defined chemical library. This could be a few 100 to several 1000 to 10,000's. Individual components from the library are added at high concentration (milimolar at least) to each well.

Application of resin and subsequent addition of the library component can be automated using commercially available robots. Following addition of the individual components to each well the plates are agitated for a pre-determined length and temperature (e.g. 30 minutes at 30°C). .

In the case of resins packed in chromatography columns, each component in the library (free of the solid support) is passed separately over the immobilized library and any eluted material collected. It is anticipated that a certain portion of the library components will be able to compete with the bound ligand and selectively liberate one or more proteins from the resins in the plates or columns. A portion of each eluate is then placed into a microtitre plate for protein analysis. Advantageously, using this method identifies the specific library component that binds to the released protein. Further analysis of the protein will determine if this binding interaction has potential therapeutic value.

If the column is broken down and distributed into the individual wells of a microtitre plate prior to releasing the bound proteome proteins, an additional step must be taken to isolate the released proteins. Library components are added to each well and incubated to release the bound proteins. Following incubation with each library component, the resin is allowed to settle or the resin suspension is centrifuged briefly (300 x g) to pellet the beeds. In accordance with one embodiment the resin beads can be magnetized beads, and a magnetic field is applied to the bottom of the plate to hold the resin at the bottom. After the resin has been separated from the supernatant, a portion of the supernatant from each well is removed and placed in a second well containing a high sensitivity protein staining reagent. This last step can be automated using standard robots familiar to those skilled in the art.

The protein detection reagent used to detect the presence of released protein can be any of those known to the skilled practitioner. In one preferred embodiment the detection reagent is one that changes color in the presence of protein. Radioactive isotopes that bind proteins (I^{125}), fluorophors (e.g. FITC) or gold stains may also be used to increase sensitivity. Specialized detection systems capable of detecting these markers are known to those skilled in the art.

Wells/column fractions that are positive for protein are selected for further analysis using standard techniques. It is anticipated that the number of

positive wells would be a few percent of the total number of components that are in the library screened – although this may be several 100 if the total number of components in the initial library numbers in the 10,000's. The proteins will be subjected to gel electrophoresis analysis, typically using SDSPAGE, to measure the purity of the protein and quantitate the amount of recovered protein. The supernatant from each well that contains liberated protein(s) is mixed with SDSPAGE sample buffer and characterized by SDS gel electrophoresis. Many hundreds of supernatants can be easily characterized using this method. The gel is stained with silver, Commassie or colloidal gold (for sequencing by mass spectrometry) or transferred to polyvinyl membrane (for mixed peptide sequencing). At this point the molecular weights and amount of protein recovered are determined.

For mixed peptide sequencing, the proteins on the polyvinyl membrane (PVM) are stained then excised (See Damer et al., (1998) *J. Biol. Chem* **273**: 24396-24405). The membrane pieces are digested briefly with CnBr, washed and placed directly into an automated Edman sequenator. Mass spectrometry can also be used but may be less desirable because of the amount of labor that is required and its inability to handle many protein samples at one time. In the case of mixed peptide sequencing between 6 and 12 rounds of Edman sequencing are carried out and the mixed peptide sequences generated sorted and matched against the databases with the FASTF (protein databases) and TFASTF algorithms (DNA data bases). This process identifies the liberated proteins in each well.

At this point a determination is made as to whether or not the liberated protein is interesting (i.e. is the protein involved in a human disease, is it an important enzyme to bacterial metabolism.. etc). If the protein is determined to have therapeutic value, then the chemical that liberated the protein from the immobilized library is chosen for further characterization using conventional approaches. For example, an affinity assay will typically be conducted to determine if the protein has sufficient affinity for the target library compound (i.e. binding at nanomolar concentrations) to be useful as therapeutic agent. In addition, analysis will be conducted to determine what is the biological impact of the interaction and whether the affinity of interaction with the targeted protein can be improved by modification.

This screen may yield several candidate proteins that are valuable in

the initial round and not confine one to a single field or market or interest. Importantly, a single round of screening will not only identify a potentially useful bioactive agent, but will also provide valuable information about its targets, what groups can be modified without affecting function and where the agents should be applied.

In another embodiment of the present invention, library components are fractionated and linked to a solid support in separate vessels (see Fig. 2). The various immobilized fractions of the library are then combined and packed into a column. The proteome of interest is then place in contact with the library under conditions suitable to allow for specific interaction between the proteome and library components. The resin is then washed to remove non-specifically bound proteins, typically using both a high ionic and a low ionic strength buffer. The bound proteins are then labeled with a detectable maker, for example either iodinating with I^{125}, or reacting with fluorescent marker or dye (e.g. iodofluorescein). The excess probe is washed away and the beads removed and individual beads placed into 96 well microtitre plates. The plates are then scanned for protein either by detecting radioactivity, fluorescence or color.

Beads that are positive for protein are treated with SDSPAGE sample buffer and their protein content determined by gel electrophoresis as described previously. If a protein is deemed to be of value, the bead that contained the identified protein is treated to liberate the ligand (library component) for chemical identification.

In some instances a library component may be identified that binds to many protein targets. As outlined in Fig. 3 a ligand that interacts with many protein targets can also be used to identify important drugs. In accordance with this embodiment a proteome is passed over a resin containing a single ligand (e.g. gamma-phosphate linked ATP Sepharose). Following washing to remove non-specifically bound proteins, either the bound proteins are labeled as described in Fig. 2 or the column is successively washed with components in a chemical library and fractions collected as described in Fig. 1. If the proteins are labeled then the beads are removed from the column and placed into microtitre plate wells (1-10 beads/well to give ~20 nmols total protein). Individual components in the library are then applied to each

well and the supernatants sampled for protein release. Preferably, each library component should be added at increasing doses in orders of magnitude ranging from 1 nM to 1 mM. Proteins that are selectively liberated in the nm-μM range are analyzed by SDSPAGE and mixed peptide sequencing is conducted as described above.

In accordance with one embodiment the method of identifying bioactive compounds from a complex mixture of proteins comprises contacting a combinatorial library with the complex mixture of proteins under conditions that allow for specific binding of the proteins to library components. Preferably the library components are immobilized on a solid support via a covalent linkage, and numerous compounds of the library are present in multiple copies that are bound to the solid support in multiple orientations. The immobilized library is then washed with a buffered solution, preferably with a high ionic strength buffer and a low ionic strength buffer, to remove non-specifically bound proteins. In accordance with one embodiment the solid support is in particulate form, and the method further comprises the step of distributing equal portions of the support particles into a plurality of wells of a microtitre plate after the step of washing the immobilized compound library.

The proteins bound to library components by specific interactions are then released, preferably by a competition reaction using one or more of the components of the library (in an unbound state). Wherein the library has been fractionated and equal portions of the support particles have been distributed into a plurality of wells of a microtitre plate the step of releasing the bound proteins comprises adding to each microtitre plate well one or more compounds of the library. The released proteins are characterized using standard techniques and the compounds of the library that specifically bind to the released proteins are identified.

The release proteins will be identified primarily based on microsequence analysis and comparison to existing protein databases. This has been made feasible because of the near completion of the nucleotide sequencing of several genomes, including human, mammalian, C.elegans, bacteria, yeast, viral, rice, corn. The invention will have increasing relevance as more species specific genomes become complete. The proteins remaining bound to the immobilized library after the washing steps can also be labeled to assist the detection of the proteins.

Isolation and identification of target proteins with an undefined immobilizedchemical library.

All of the initial steps are the same as described above. Following preparation of the resins and application of protein mixture, the resins are again aliquoted into titre plate wells. However, since the components in the library are poorly defined and of unknown number some fractionation of the library using standard methods is required. Useful steps for fractionation include organic and aqueous extraction, HPLC or ionic-exchange fractionation. The fractionated library is then applied to each well containing resin and incubated as described. In one embodiment these steps are carried out robotically. The column chromatography approach out-lined above is also applicable with an undefined library.

Following incubation with each fraction of the library the resin is pelleted as described above and a sample of the supernatant taken for protein analysis. Fractions that contain liberated protein are selected for characterization by SDSPAGE. At this point it is likely that some fractions of the library will liberate many proteins, some only a few. In either case, mixed peptide sequencing or mass spectrometry can be used to identify all these components in a short space of time. With mixed peptide sequencing a standard Edman sequencer containing 4 reaction chambers can identify 20-30 proteins per week. Mass spectrometry will be somewhat slower if a species-specific database is not available. The list of proteins that are identified for each well is then surveyed for the criteria stated earlier.

Proteins that are deemed valuable are expressed as recombinant proteins and immobilized on a second resin (Sepharose or agarose beads). The fraction of interest or entire chemical library is then passed over the protein affinity column to selectively recover the chemical compounds with in the library that bind the protein of interest. Mass spectrometry or NMR techniques can then be used to identify or determine the structure of the bioactive compound. One then proceeds with the standard methods to characterize the bioactivity of the isolated chemical. All three strategies as outlined in Figs. 1-3 can be applied to an undefined library.

In accordance with an alternative embodiment bioactive compounds are isolated through the use of intact cells. This method is particularly useful for identifying bioactive agent that interact with cell surface molecules such as receptors.

The cells are grown on a cell culture substrate suitable for the type of cell grown. A complex mixture of labeled proteins or peptides is then added to the cells under conditions that allow for specific binding of the labeled proteins to the cell surface proteins. In one embodiment the cells are culture in multiple plates and the complex mixture of labeled proteins/peptides is divide equally among the multiple plate of cells. The cells are then washed under conditions that do not dislodge the cells from the substrate. In accordance with one embodiment the cells can be fixed to the cell culture substrate prior to incubation with the labeled proteins.

After the plates have been washed to remove non-specifically bound proteins, the plates are screened for the presence of labeled proteins. The labeled proteins/peptides are released using the same procedures as described above and analyzed by gel electrophoresis and microsequencing.

In accordance with one embodiment, the complex mixture of labeled proteins comprises randomly generated peptide sequences that have been labeled with a fluorescent entity. The binding of such labeled proteins to the cell surface molecules effectively concentrates the label at the bottom of the culture plate and thus a positive reaction can be detected even in the absence of washing the cells to remove unbound labeled protein. For example, an excitation light source can be provided wherein the beam of light is parallel to the cell surface and the detector is similarly position so that only signal generated from the cell surface will be detected. The bound protein can then be released using any of the techniques described previously herein, and the protein can be analyzed as describe above.

Fig. 4 exemplifies one approach used in accordance with the present invention for identifying cell surface receptors and their peptide ligands en masse. The overall scheme outlined in Fig. 4 is a variation of that disclosed in Fig. 2. Bioactive peptides are of pharmaceutical value because they mimic naturally occurring proteins or larger peptides that bind to important cell surface receptors e.g. interferon's, cytokines, growth factors, endorphins. Bioactive peptides can be generated randomly in large libraries using combinatorial approaches. Peptides in these libraries are general range from 4 to 20 amino acids in length. These libraries can be generated synthetically or recombinantly as fusion proteins. In the case of fusion proteins, random peptide sequences are displayed at the N or C termini of a

known recombinant protein expressed in yeast or bacteria (Blum et al. 2000 *PNAS* 97, 2241; Geyer et al. 1999 *PNAS* 96: 8567). The fusion protein displaying the peptide are recovered by affinity chromatography through an affinity tag that is present at the opposite end of the molecule (C or N termini) from the peptide.

There are 24 naturally occurring amino acids in nature which can be used to construct a random combinatorial library. A peptide library consisting of peptides of 20 amino acids in length can therefore have 20^{24} possible combinations. This gives an extremely large number of possible variations of peptide and theoretically covers all possible peptide sequences that could occur in nature. Typically these types of peptide libraries can be used to search for cell surface receptors or protein partners that would selectively bind one or more peptide sequences contained within them. As shown in Fig. 4, a recombinant peptide library cultured in bacteria or yeast, or a synthetic combinatorial peptide library tagged with a fluor (e.g. fluorescein), is mixed (at 1nM - 1μM concentration) with a designated target cell line (e.g. cancer cell line, B or T cell) that is present in the wells of a multi chamber titer plate. The plate is placed in an instrument capable of detecting fluorescent labeled probes on cell surfaces at 100 -5000 molecules per cell. In our laboratory we use the PE-Biosystems FMAT robot (Swartzman et al. 1999 Anal. Biochem. 271: 143). In the case of synthetic peptides the cells are screened for specific binding of fluor tagged peptides on the cell surface. In the case of recombinant fusion proteins displaying the random peptides a fluor tagged antibody that recognizes the fusion protein is added.

The peptides that produce positive results are sequenced. In the case of the synthetic peptides this can be done directly without further purification in an Edman sequencer or mass spectrometer. In the case of the bacterially or yeast expressed fusion protein two methods can be used to sequence the peptide. Positive clones can be cultured and the expression vector encoding the fusion protein can be sequenced across the region encoding the random peptide by DNA sequencing. Or alternatively, the fusion protein can be isolated from a culture of bacteria or yeast and the random peptide sequence determined by mass spectrometry or Edman sequencing. Once the peptide sequence is identified it is produced synthetically and tagged with a fluorophor. The affinity of the peptide for the cell surface receptor is determined. If

the affinity is determined to be sub μM, an affinity column is constructed from the peptide for purification of the receptor target. Typically this would involve linking the peptide via its C or N terminus to a C8 spacer attached to a Sepharose bead. Cell extract prepared from the designated cell target would then be passed over the resin to recover the receptor target for identification by protein sequencing. Bound proteins would be recovered by either eluting the resin with free peptide or washing with SDS.

SDSPAGE and protein staining would be used to quantitate and evaluate purity. Mixed peptide sequencing or mass spectrometry would be used to identify the protein. If the protein is found to be an important cell surface receptor of commercial or medical value the peptide and protein target would be fully characterized. This screen is anticipated to identify many cell surface receptors and their bioactive peptide ligands. Some receptors will be well characterized, many others are anticipated to be novel. Significantly, in addition to identifying new peptide ligands and their physiological targets, our assay, like the other methods, also gives a measure of selectivity and potential toxicity. This is because the identified bioactive ligands, in addition to their true physiological target, had an equal opportunity to interact with all other cell surface receptors that happen to be expressed on the designated cell target.

Example 1

Isolation of Adenine Nucleotide Binding Protein from a Proteome

As a proof of principle and to evaluate the types of proteins that bind to gamma-phosphate linked ATP-Sepharose, tissue extracts prepared from rabbit skeletal muscle, liver, kidney, brain or bladder were passed over a gamma-phosphate linked ATP-Sepharose affinity resin. Following extensive washing to remove non-specifically associated proteins, the resin was washed sequential with NADH, AMP, ADP and ATP. Fig. 5 shows the results from characterization of proteins isolated from skeletal muscle. Similar results were obtained with other tissues, although the pattern, abundance and complexity between tissues varied considerably due to varied levels of expression of individual proteins (See Fig. 5 and Table 1).

Gamma phosphate linked ATP -Sepharose was washed with extracts prepared from rabbit, skeletal muscle, kidney, liver, brain or bladder. Following

washing the resin was eluted successively with the indicated nucleotides as described in Fig. 5.

Proteins 1-17 (see Fig. 5) were identified by mixed peptide sequencing (Table 1). Eluted proteins were characterized by SDSPAGE, transferred to PVM and treated with CnBr or Skatol prior to mixed peptide sequencing. Mixed peptide sequencing was carried out on average for 6-12 Edman cycles. The mixed sequences were sorted and matched against the entire published protein or DNA data bases with the FASTF or TFASF algorithms respectively (Damer et al. 1998., Alms et al. 1999). Expectation scores for the identified proteins ranged from $2.6\ e^{-7}$ for PKA to $1.2\ e^{-54}$ for GAPDH. Expectation scores after each search for the next highest scoring non-related protein were generally $< 2.3\ e^{-4}$ The experiment shown was repeated on several occasions, and on several different tissue including liver (120g) , kidney (60g), brain (60g) and bladder (20g) with similar results (Table 1).

Table 1			
NADH	pmol	AMP	pmol
GAPDH (M) 1*	25,000.0	Purine synthetase (M,L) 2*	5.0
malate dehydrogenase (M) 8*	0.5	Phosphorylase (M) 3*	10.0
glutamate dehydrogenase (L) 9*	1.5	AMP activated protein kinase (L) #	2.0
aldehyde dehydrogenase (M) 7*	2.0	EST AA254816 (L)	4.0
lactate dehydrogenase (M,L) #	50.0	ESTAA571903 (L)	5.0
6-phosphogluconate dehydrogenase (M,L)	5.0	phosphatidylinositol-4-phosphate 5-kinase (L)	2.0
isocitrate dehydrogenase (L)	1.0	protein kinase DUN1 (related) (L)	1.0
3-hydroxyacyl-CoA dehydrogenase (L)	1.0		
sorbitol dehydrogenase (L)	2.0		
alcohol dehydrogenase (M,L)	20.0		
glucose-6-phosphate dehydrogenase (M,L)	17.0		

Table 1 (cont.)			
ADP	pmol	ATP	pmol
Heat shock protein 90 (M/L) 4*#	15,000.0	MAPK (M) 5*#	5.0
Purine synthetase (M/L) 2*	14.0	MEK1 (M)#	6.0
		Pyridoxal kinase (M, L) 12*	10.0
		Arginnosuccinate synthetase (M) 11*	1.0
		Glutamate ammonium ligase (M,L) 6*	2.0
		Adenosine kinase (M,L) 15*	5.0
		CSK (M) 16*#	4.0
		HSPA5 (M) 17*	4.0
		P90 S6 kinase (M, L) 14*	0.5
		P70 S6 kinase (L)	0.2
		Pyridoxal kinase (L)	5.0
		P98 glucose induced kinase (M,L,SM,B,K) 10*	10.0
		Heat shock protein 70 (L,M,B,SM,K)	55.0
		PKA (L,M,B,SM)	6.0
		Glutamine synthetase (L)	9.0
		ZIP Kinase (SM)#	0.2
		Phosphofructokinase (L)	6.0
		Heat shock protein 60 (M,L,B,SM,K)	10.0
		RNA helicase (L)	1.0
		Protein kinase PC-1 (L)	0.6
		Protein kinase C epislon (L)#	0.2
		Beta tubulin (B)#	50

		P6 electron transport flavoprotein α subunit (L)	1.0
		serine/theonine-protein kinase ip11 (related) (L)	0.2
		protein kinase C beta-II (L)	0.2
		protein kinase kem (L)	0.3
		cyclic G kinase (SM)#	5.0
		lupus nephritis protein LN1 (SM)	1.0
		casein kinase 1 (M)	5.0
		casein kinase 11 (M)	6.0
		GSKIII (M)	4.0
		lim domain kinase 1 (SM)	0.2
		protein kinase pkx1 (SM)	0.3
		pp60c-src (M)	1.0
		MLCK (M,SM)#	6.0
		Phosphorylase kinase (M)	20.0
		Arginine deimidase (M) 18*	5.0
		CAM kinase II (B,M)#	1.0
		fructose-1-6-bisphosphatase (M,L)	2.0

The pmol amount of protein recovered was determined from PTH amino acid recovery during mixed sequencing multiplied by the volume applied to the gel and fraction volume (mls). *Indicates proteins identified in Fig. 5; # indicates proteins tested for binding efficiency by Western analysis of the column flow through.

To unambiguously identify the eluted proteins in each case, peak column fractions were transferred to PVM following SDSPAGE and subjected to mixed peptide sequencing (Table 1). Table 1 shows the identification of over 70 proteins that were eluted from ATP Sepharose loaded with skeletal muscle, liver, brain, kidney or bladder. Without exception, all of the proteins identified in the protein data bases belonged either to the protein kinase, dehydrogenase or ATP-grasp

classes of purine binding proteins. Analysis of PTH amino acid recovery during mixed peptide sequencing reveals that the affinity resin recovered proteins of both high and low cell copy number (Table 1). Western blotting of the column flow through with antibodies to several of the identified proteins demonstrated that the resin absorbed the tested proteins with >85 % efficiency from the initial extract (Table 1). This finding indicated that the differences in recovery of individual proteins in the nucleotide washes was a reflection of cell copy number rather than because of affinity differences for the immobilized ATP.

Several of the identified proteins have been crystallized with NADH, AMP, ADP or ATP bound and these published structures explain selective recovery of each classification of protein from the affinity resin. Inspection of the three dimensional structure of 10 of the dehydrogenases identified in the NAD wash shows in each case the adenine portion of these nucleotides is buried within a cleft containing the conserved Gly-X-Gly-X-X-Gly loop. The diphosphate portion of the bound nucleotides spans an open region on the surface connecting to the nicotinamide moiety accommodated within a closely situated second binding site.

The finding that 0.5 mM NADH/NAD exclusively eluted dehydrogenases over other purine binding proteins is consistent with the well established preferences of these types of enzymes for nicotinamide containing purines. Although it should be noted that in separate experiments increasing the concentration of NADH/NAD to >10 mM did begin to elute many of the proteins found in the subsequent AMP wash. Characterization by microsequencing of the AMP eluate identified two proteins that are allosterically regulated by the nucleotide, glycogen phosphorylase and the AMP activated protein kinase. In the case of phosphorylase, elution with AMP is consistent with the crystal structure of enzyme in its 'T' state, which is favored by the presence of glucose, ADP and ATP competitors, low concentrations of substrate (Sprang et al. 1991). Purification of the AMP activated protein kinase over gamma phosphate linked ATP-Sepharose has been reported previously (Davis et al 1996). The enzyme is known to contain both an ATP and AMP binding pocket and is activated by AMP *in vitro*. Recovery of the kinase with AMP therefore is most likely because of interaction with the immobilized ATP with its AMP binding pocket. Elution of multifunctional protein ADE2 with AMP (and

also ADP) is consistent with involvement of this protein in catalyzing the conversion of AIR to CAIR (steps 6 and 7) in purine biosynthesis. Although the three dimensional structure of mammalian ADE2 has not been solved, the related *E.coli* enzyme, N5-carboxyaminoimidazole ribonucleotide synthetase (PurK) with ADP bound was recently reported (Thoden et al. 1999). PurK belongs to the ATP grasp superfamily of purine binding proteins, and in prokaryotes, plants and yeast, catalyzes the conversion of AIR to CAIR in a two step process (steps 6 and 7 of 10) involving a second distinct gene product PurE. Recovery of mammalian ADE2 from the affinity resin in this present study suggests it also binds purine nucleotides in a similar orientation to that found in PurK. Phosphatidylinositol-4-phosphate 5-kinase, protein kinase DUN1 (related) and the two EST AA254816, ESTAA571903 all contain nucleotide binding motifs in their primary sequence and elution of these proteins with AMP suggests that they also bind the purine orientating the phosphate such that it is solvent accessible.

Elution of the resin with ADP following AMP eluted two proteins, HSP90 and ADE2 in all tissues tested. The recovery of additional ADE2 with ADP suggests that this enzyme may either have two separate adenine binding pockets or exist in two conformational states that discriminate the presence of _ and _ phosphates on the two types of purine. Recovery of HSP90 with ADP is consistent with recent reports by Toft and co-workers identifying the N terminal domain of HSP 90 as a Mg 2+ ATP/ADP binding domain and the crystal structure of this domain with ADP or ATP bound (Prodromou et al. 1997, Stebbins et al. 1997 35-38). Interestingly, the purine binding pocket on HSP90 was not readily predicted to exist based on primary sequence alignments alone. Recovery of HSP90 suggests that other non-conventional purine binding proteins presenting adenine containing nucleotides in the "protein kinase" orientation are likely to be recovered from the affinity resin. Examples of other proteins that been crystallized with purine bound and classified as having non-conventional binding domains are the adenine binding domain of DNA gyrase B (Wigleyet al. 1991), the AMP binding sites on glycogen phosphorylase and adenylate kinase, the ADP binding sites on fructose-1-6-bisphosphatase and phosphofructokinase, the cyclic AMP binding sites on catabolite activator protein and ATP binding site in DD-ligase. Notably four of these proteins were subsequently

recovered in the ATP wash (Table 1).

Final elution of the affinity resin with ATP eluted a diverse range of proteins in all tissues tested, from heat shock proteins and metabolic enzymes with non-conventional nucleotide binding folds to a variety of protein kinase family members (Table 1). The majority of the proteins recovered are known to utilize Mg 2+ ATP and show a high degree of specificity towards the nucleotide. Consistent with this observation, we have found that inclusion of mM NADH, AMP and ADP in the extraction buffer completely abolishes binding of all of the proteins shown in Table 1 that would normally be recovered from the resin in their absence (data not shown). In contrast, proteins identified in the ATP elution are retained on the resin under these conditions. Amongst the most frequent of all the adenine binding proteins sequenced in Table 1 are protein kinases. It has been estimated that up to 2% of the entire human genome may encode a protein kinase with a highly conserved ATP binding cassette. When the amino acid sequences of over 400 protein kinases are aligned with that of cyclic AMP dependent protein kinase, 15 amino acids residues within 11 conserved subdomains are nearly invariant. In addition, there are 19 hydrophobic amino acids of similar structure that are also conserved within the protein kinase family. In the activated state, the conserved and invariant amino acids of the ATP binding cassette make intimate contact with MgATP and orientate the molecule such that its gamma phosphate is exposed at the lips of the catalytic cleft. The recovery of several distinct tyrosine and serine/threonine protein kinases as reported herein, and reports by others utilizing gamma phosphate linked ATP-Sepharose in purification schemes for specific kinases, demonstrates that this affinity resin is likely to bind all protein kinase family members. Furthermore, this finding, combined with the frequency of occurrence of protein kinases, dehydrogenases and some of the other proteins identified in Table 1 in the current protein and DNA data bases suggests that the ATP resin may catch 3-5% of all proteins present in most eukaryotic genomes.

Example 2

Screening for Selective Inhibitors for Purine Binding Proteins

To test the concept of proteome mining of a combinitorial and natural products small molecule libraries for selective inhibitors of purine binding proteins, geldanamycin (GA) and 74 structural analogs were passed over gamma-phosphate linked ATP Sepharose that had previously been loaded with whole skeletal muscle extract. To ensure that all proteins that bind adenosine in the "protein kinase orientation" the resin contained an ATP concentration between 10-15 μMol/ml. Initial ligand screens were performed at 10 μM which would enable only pharmacologically relevant competitive inhibitors to be identified in the small library. This is because any protein that was selectively eluted from the ATP resin, by a particular GA analog, in order to have pharmacological relevance in subsequent cell based assays would have to be able to compete with a physiological ATP concentration of ~10 mM. As discussed earlier, the high ligand concentration also ensured that proteins of both high and low affinity, and copy number would be equally and maximally recovered by the resin. Fig. 6 shows a silver stain of peak column fractions after eluting the affinity resin with increasing concentrations of GA. A similar gel was transferred to PVM and the most abundant proteins identified by mixed peptide sequencing.

Washing the affinity column with 10 nM GA was found to almost exclusively elute a single protein at 45kDa. The fraction also contained a minor amount of a 90 kDa protein. Mixed peptide sequencing identified the 45kDa protein as ADE2 and the 90kDa protein as HSP90. In particular, mixed peptide sequencing identified peptide sequences Met Phe Phe Lys Asp Asp Ala Asn Asn Asp Pro Gln Trp (SEQ ID NO: 1) and Met Lys Ile Glu Phe Gly Val Asp Val Thr Thr Lys Glu (SEQ ID NO: 2) which were 100% identical to peptide sequences of purine multifunctional enzyme (ADE2). Mixed peptide sequencing also identified peptide sequences Met Thr Lys Ala Asp Leu Ile Asn Asn (SEQ ID NO: 3) and Met Ile Gly Gln Phe Gly Val Gly Phe Tyr (SEQ ID NO: 4) which were 100% identical to peptide sequences of heat shock protein 90 beta (HSP90).

These findings identify ADE2 as a new target for GA *in vitro*. PTH analysis of the sequenced proteins indicated that there were 2 pmols of ADE2 in the

gel compared with <150 fmol of HSP90. Increasing the concentration of GA to 100 nM eluted 2 proteins of 70kDa and 65 kDa respectively that were identified by mixed peptide sequencing as proteolytic fragments of HSP90. No other proteins were eluted from the resin until the concentration of GA reached 10 μM. Approximately 1pmol of HSP90 was recovered in the gel at this step. Increasing the concentration of GA to 100 μM eluted a large amount of HSP90 and several N terminal proteolytic fragments of the protein. Approximately 11 pmols of the proteins was sequenced from the gel.

At 1mM GA 3 pmols of HSP90 was sequenced from the gel. Significantly no other proteins were eluted at this concentration. Subsequent, elution of the column with 10 mM ATP liberated a complex mixture of proteins of varying abundance and molecular weights. Consistent with previous results, mixed peptide sequencing of a selection of these proteins identified them as adenine nucleotide binding proteins.

The finding that the GA concentration could be raised by 3 orders magnitude over the concentration required to elute ADE2 before significant HSP90 was recovered demonstrates that in solution GA prefers the former enzyme over the later. Although, increasing the concentration to 1 mM did not elute any other proteins and is a testament to the selective of GA towards HSP90 and ADE2, these findings have implications for the actions of GA *in vivo*. In particular, the elution of ADE2 by GA illustrates a potential unforeseen toxicity of GA. The enzyme ADE2 catalyzes two essential steps that are required for the synthesis of purine nucleotides. Inhibition of ADE2 activity would therefore be toxic to all cell types.

To discriminate functional regions on GA that may discriminate sites of interaction with HSP90 from ADE2, 74 structural analogs (Table 2) of the molecule were passed over gamma-phosphate linked ATP Sepharose that had been charged with skeletal muscle extract.

Each analog of GA was washed over the gamma-phosphate linked ATP Sepharose that had been charged with skeletal muscle extract at 10 μM and the elutes analyzed by fluorography and SDSPAGE following tagging of the eluted proteins with iodofluorescein. The eluted proteins were visualized by laser induced fluorescence using a molecular dynamics flat bed imager. Using this method proteins that contained at least one reduced cysteine residue could be detected in the eluate at

<0.1 fmol. Following fluorpohor labeling the eluted proteins were characterized by SDSPAGE and fluorography. Mixed peptide sequencing identified the eluted proteins as either HSP90 or ADE2. Several compounds within the small GA analog library have selectivity for HSP 90, while others are more selective for ADE2.

Assay of purified rabbit and human ADE2 confirmed that all compounds that selectively eluted ADE2 from the affinity also inhibited the enzyme in vitro (Table 3). All assays were performed against purified human ADE2. Results shown are from three separate experiments. Compounds that eluted HSP90 selectively had no effect of ADE2 activity in vitro at μM concentrations. Significantly, compounds that selectively eluted HSP90 showed low biological activity in cell based growth inhibition assays. In contrast, compounds that showed selectivity for ADE2 were potent inhibitors of cell growth. These findings demonstrate that *in vivo*, the biological effects of geldanamycin are because of its ability to inhibit ADE2 activity rather than through any actions on HSP90.

Table 3. Determination of apparent Ki for ADE2 against GA structural analogs

Inhibitor	K_i^{-1}(μM)	Inhibitor	K_i^{-1}(μM)
210760	6.803	683663	37.037
189794	3.546	330506	45.455
182857	5.525	255109	38.462
189795	0.688	320877	30.303
255111	5.051	665479	2.037
189793	3.497	265482	10.417
604169	38.462	320947	21.277
697886	15.152	182858	58.824
683662	16.667	156219	1000.000
672165	1.250	683660	90.909
330500	500.000	320946	33.333
661581	6.250	330509	21.277
255107	62.500	156217	22.222
683201	11.905	169627	7.634
655480	7.143	156218	58.824
255104	23.810	359658	1.815
682299	13.514	658515	0.484
655746	0.688	236651	34.483
682300	18.868	651937	0.185
330510	12.658	330499	0.615
683666	90.909	683664	1.692
661580	14.493	707545	62.500
662199	58.824	210753	0.362
690214	17.544	662199	50.000
607306	0.792	320944	20.000
607307	1.629	236652	10.526
255110	22.222	48810	14.706
321593	20.833	156216	26.316
674124	13.699	19990	20.408
156215	45.455	210761	27.778

Claims:

1. A method of identifying bioactive compounds, said method comprising the steps of

contacting a combinatorial library with the protein members of a proteome, to bind said protein members to compounds of said library, wherein the compounds of said library are immobilized on a solid support;

washing the library with a buffered solution;

releasing the bound proteins;

characterizing the released proteins; and

identifying the compounds of the library that bind to the released proteins.

2. The method of claim 1 wherein the step of releasing the bound proteins comprises contacting the library with one or more compounds of said library.

3. The method of claim 1 wherein the step of washing the library comprises the steps of washing with a high ionic strength buffer and a low ionic strength buffer.

4. The method of claim 1 wherein the compounds of said library are covalently bound to said solid support.

5. The method of claim 4 wherein each of the library compounds are present in multiple copies that are bound to the solid support in multiple orientations.

6. The method of claim 5 wherein the solid support is in particulate form, and the method further comprises the step of distributing equal portions of the support particles into a plurality of wells of a microtitre plate after the step of washing the immobilized compound library.

7. The method of claim 6 wherein the step of releasing the bound proteins comprises adding to each microtitre plate well one or more compounds of the library.

8. The method of claim 1 further comprising the step of labeling the bound proteins after the washing step.

9. The method of claim 8 wherein the immobilized compounds are bound to polymer beads and the method further comprising the step of distributing single beads into separate wells of a microtitre plate, after the step of labeling the bound proteins.

10. The method of claim 9 wherein the step of releasing the bound proteins comprises contacting the individual beads with a chaotropic agent.

11. A method of identifying bioactive compounds present in a proteome, said method comprising the steps of

contacting an immobilized ligand with the protein members of a proteome;

washing the immobilized ligand with a buffered solution;

contacting the bound proteome proteins with a target compound to release bound proteome proteins;

collecting the released bound proteins; and

determining the identity of the released proteins.

12. The method of claim 11 wherein the target compound is the same as the immobilized ligand.

13. The method of claim 11 wherein the target compound is a functional analog of the immobilized ligand.

14. The method of claim 11 further comprising the step of labeling the bound proteins before the step of contacting the immobilized compound library with the individual component of the compound library.

15. A method of identifying bioactive compounds from a poorly defined immobilized combinatory library, said method comprising the steps of

contacting a the c combinatory library with the protein members of a proteome, to bind said protein members to compounds of said library, wherein the compounds of said library are immobilized on a solid support;

washing the library with a buffered solution;

releasing the bound proteins;

immobilizing a released protein on a solid support;

contacting the immobilized released protein with compounds of said combinatorial library to bind components of the combinatorial library to the immobilized released protein;

washing the immobilized released protein with a buffered solution;

releasing the bound library components; and

identifying the released library components that bind to the immobilized released protein.

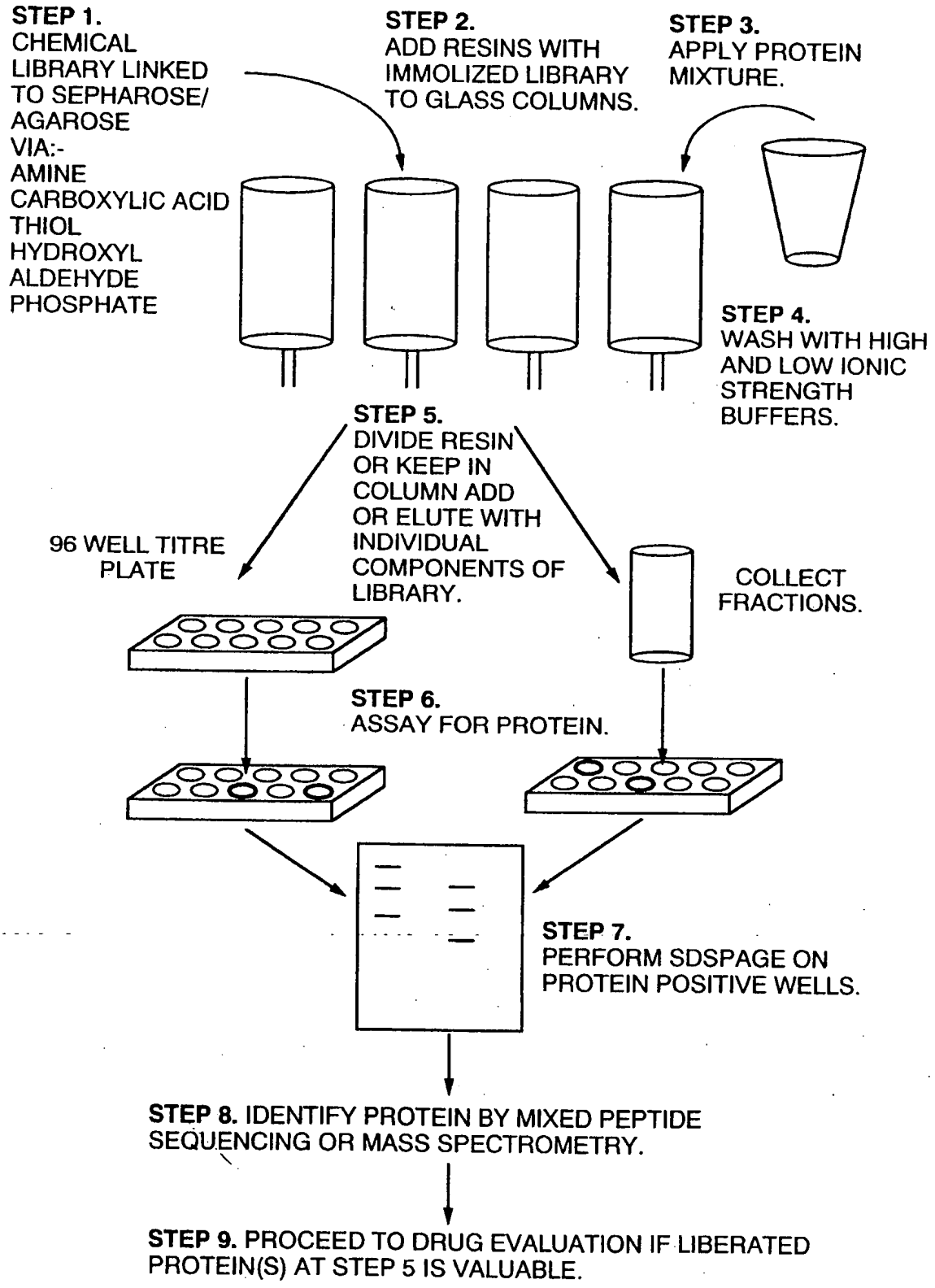

STEP 1.
CHEMICAL LIBRARY LINKED TO SEPHAROSE/ AGAROSE VIA:-
AMINE
CARBOXYLIC ACID
THIOL
HYDROXYL
ALDEHYDE
PHOSPHATE
STEP 2.
ADD RESINS WITH IMMOLIZED LIBRARY TO GLASS COLUMNS.
STEP 3.
APPLY PROTEIN MIXTURE.
STEP 4.
WASH WITH HIGH AND LOW IONIC STRENGTH BUFFERS.
STEP 5.
DIVIDE RESIN OR KEEP IN COLUMN ADD OR ELUTE WITH INDIVIDUAL COMPONENTS OF LIBRARY.
96 WELL TITRE PLATE
COLLECT FRACTIONS.
STEP 6.
ASSAY FOR PROTEIN.
STEP 7.
PERFORM SDSPAGE ON PROTEIN POSITIVE WELLS.
STEP 8. IDENTIFY PROTEIN BY MIXED PEPTIDE SEQUENCING OR MASS SPECTROMETRY.
STEP 9. PROCEED TO DRUG EVALUATION IF LIBERATED PROTEIN(S) AT STEP 5 IS VALUABLE.

FIG. 1

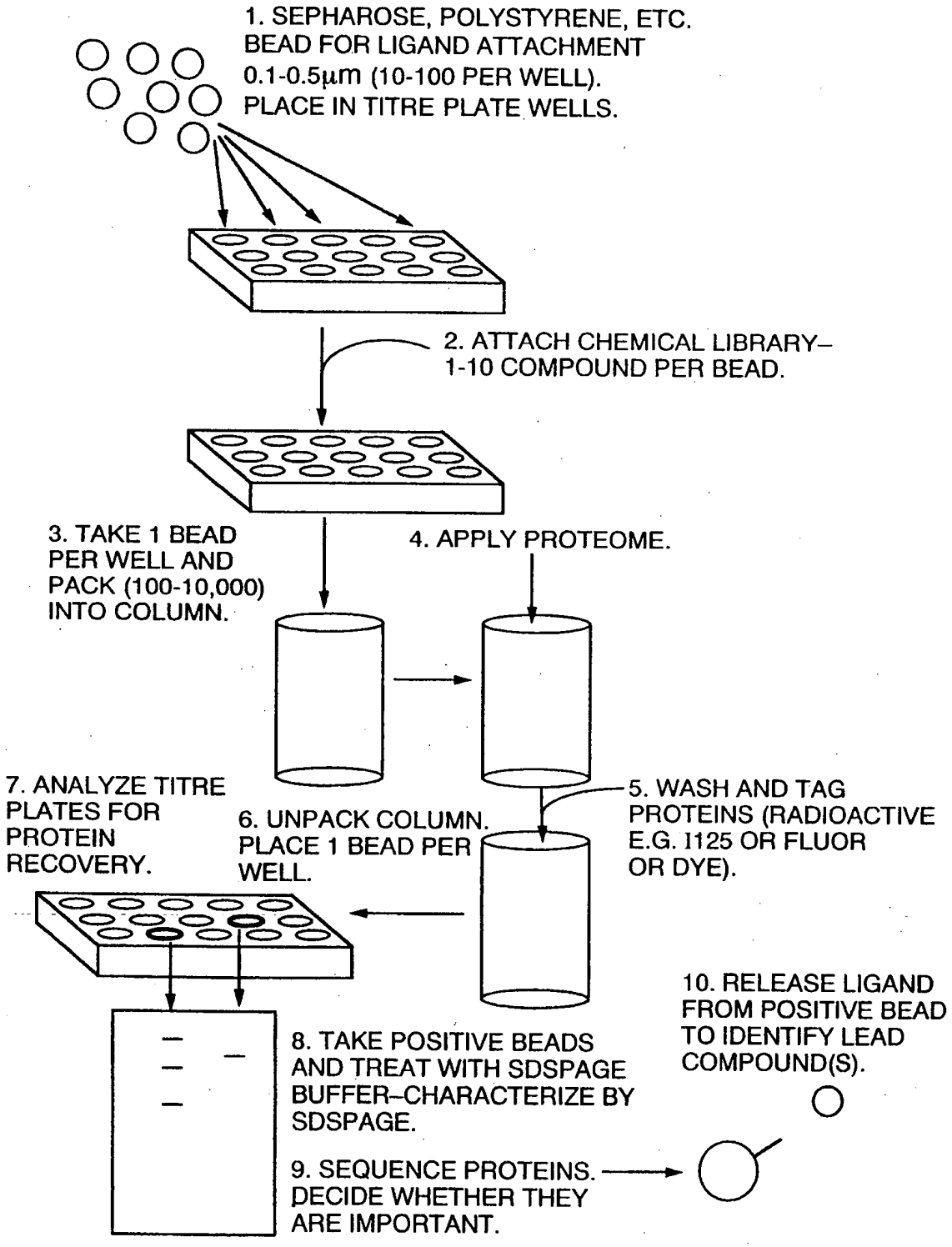

1. SEPHAROSE, POLYSTYRENE, ETC. BEAD FOR LIGAND ATTACHMENT 0.1-0.5μm (10-100 PER WELL). PLACE IN TITRE PLATE WELLS.
2. ATTACH CHEMICAL LIBRARY– 1-10 COMPOUND PER BEAD.
3. TAKE 1 BEAD PER WELL AND PACK (100-10,000) INTO COLUMN.
4. APPLY PROTEOME.
5. WASH AND TAG PROTEINS (RADIOACTIVE E.G. I125 OR FLUOR OR DYE).
6. UNPACK COLUMN. PLACE 1 BEAD PER WELL.
7. ANALYZE TITRE PLATES FOR PROTEIN RECOVERY.
8. TAKE POSITIVE BEADS AND TREAT WITH SDSPAGE BUFFER–CHARACTERIZE BY SDSPAGE.
9. SEQUENCE PROTEINS. DECIDE WHETHER THEY ARE IMPORTANT.
10. RELEASE LIGAND FROM POSITIVE BEAD TO IDENTIFY LEAD COMPOUND(S).

FIG. 2

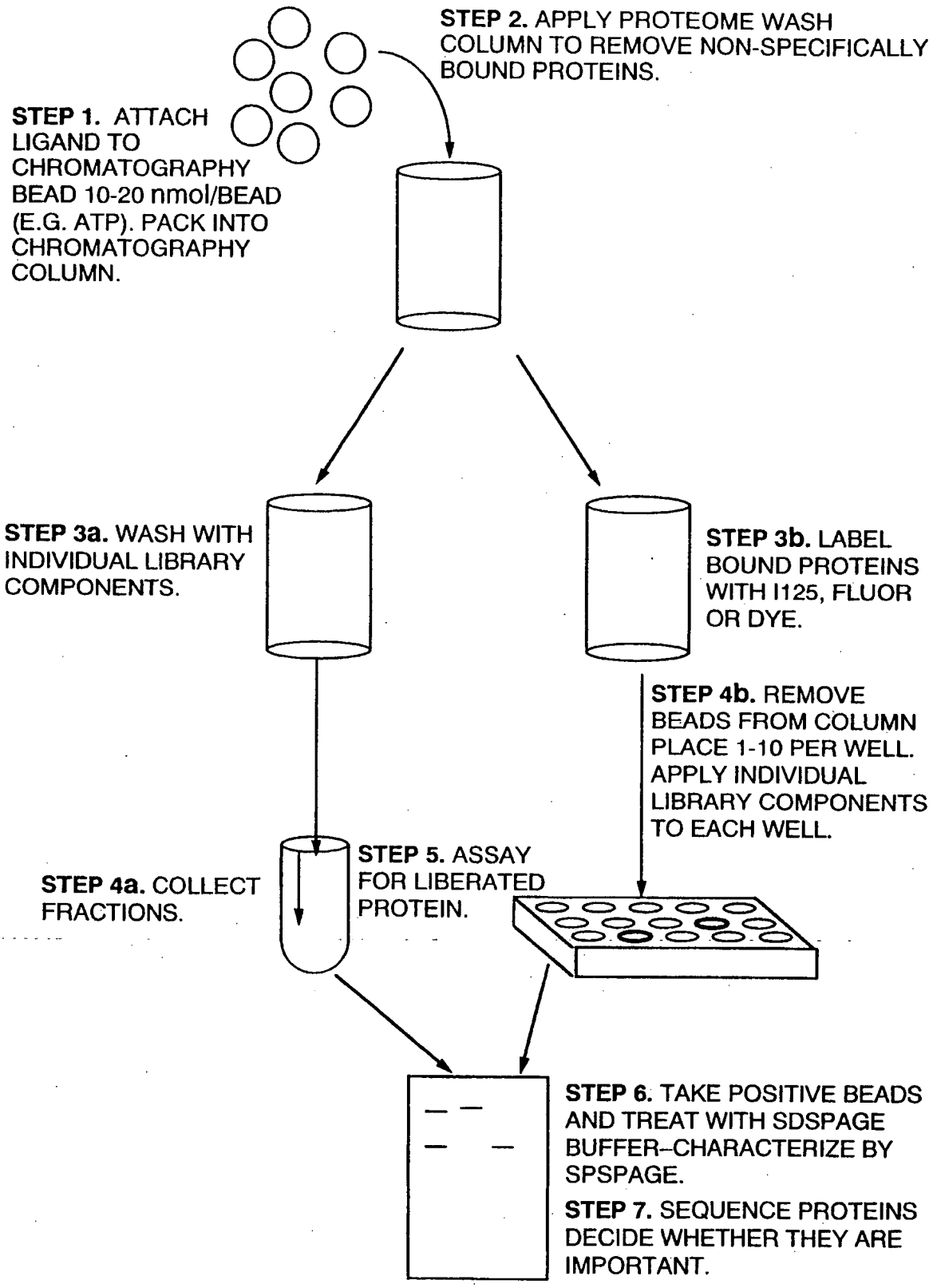

STEP 1. ATTACH LIGAND TO CHROMATOGRAPHY BEAD 10-20 nmol/BEAD (E.G. ATP). PACK INTO CHROMATOGRAPHY COLUMN.
STEP 2. APPLY PROTEOME WASH COLUMN TO REMOVE NON-SPECIFICALLY BOUND PROTEINS.
STEP 3a. WASH WITH INDIVIDUAL LIBRARY COMPONENTS.
STEP 3b. LABEL BOUND PROTEINS WITH I125, FLUOR OR DYE.
STEP 4b. REMOVE BEADS FROM COLUMN PLACE 1-10 PER WELL. APPLY INDIVIDUAL LIBRARY COMPONENTS TO EACH WELL.
STEP 4a. COLLECT FRACTIONS.
STEP 5. ASSAY FOR LIBERATED PROTEIN.
STEP 6. TAKE POSITIVE BEADS AND TREAT WITH SDSPAGE BUFFER--CHARACTERIZE BY SPSPAGE.
STEP 7. SEQUENCE PROTEINS DECIDE WHETHER THEY ARE IMPORTANT.

FIG. 3

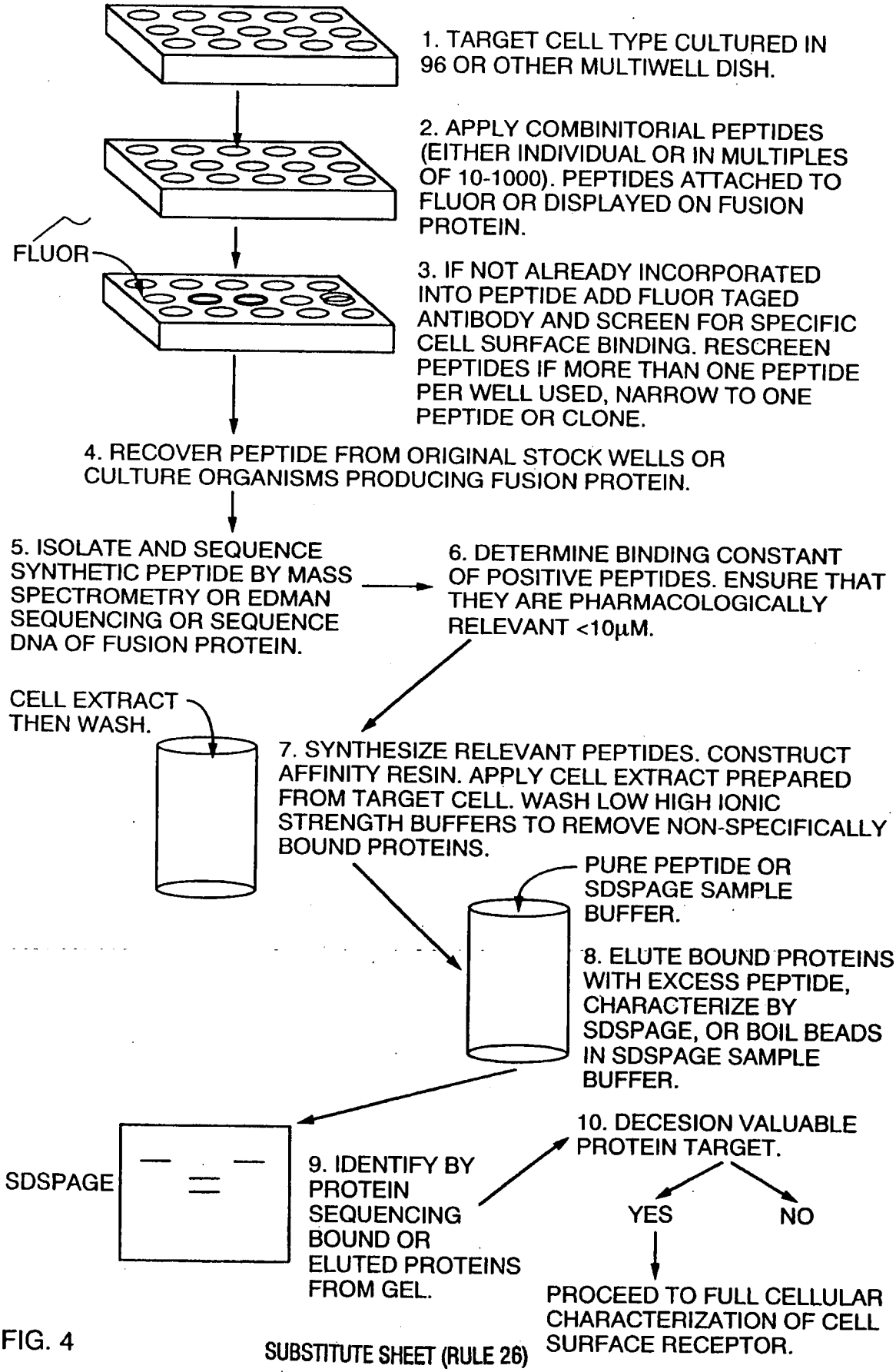

1. TARGET CELL TYPE CULTURED IN 96 OR OTHER MULTIWELL DISH.
2. APPLY COMBINITORIAL PEPTIDES (EITHER INDIVIDUAL OR IN MULTIPLES OF 10-1000). PEPTIDES ATTACHED TO FLUOR OR DISPLAYED ON FUSION PROTEIN.
FLUOR
3. IF NOT ALREADY INCORPORATED INTO PEPTIDE ADD FLUOR TAGED ANTIBODY AND SCREEN FOR SPECIFIC CELL SURFACE BINDING. RESCREEN PEPTIDES IF MORE THAN ONE PEPTIDE PER WELL USED, NARROW TO ONE PEPTIDE OR CLONE.
4. RECOVER PEPTIDE FROM ORIGINAL STOCK WELLS OR CULTURE ORGANISMS PRODUCING FUSION PROTEIN.
5. ISOLATE AND SEQUENCE SYNTHETIC PEPTIDE BY MASS SPECTROMETRY OR EDMAN SEQUENCING OR SEQUENCE DNA OF FUSION PROTEIN.
6. DETERMINE BINDING CONSTANT OF POSITIVE PEPTIDES. ENSURE THAT THEY ARE PHARMACOLOGICALLY RELEVANT <10μM.
CELL EXTRACT THEN WASH.
7. SYNTHESIZE RELEVANT PEPTIDES. CONSTRUCT AFFINITY RESIN. APPLY CELL EXTRACT PREPARED FROM TARGET CELL. WASH LOW HIGH IONIC STRENGTH BUFFERS TO REMOVE NON-SPECIFICALLY BOUND PROTEINS.
PURE PEPTIDE OR SDSPAGE SAMPLE BUFFER.
8. ELUTE BOUND PROTEINS WITH EXCESS PEPTIDE, CHARACTERIZE BY SDSPAGE, OR BOIL BEADS IN SDSPAGE SAMPLE BUFFER.
SDSPAGE
9. IDENTIFY BY PROTEIN SEQUENCING BOUND OR ELUTED PROTEINS FROM GEL.
10. DECESION VALUABLE PROTEIN TARGET.
YES
NO
PROCEED TO FULL CELLULAR CHARACTERIZATION OF CELL SURFACE RECEPTOR.

FIG. 4

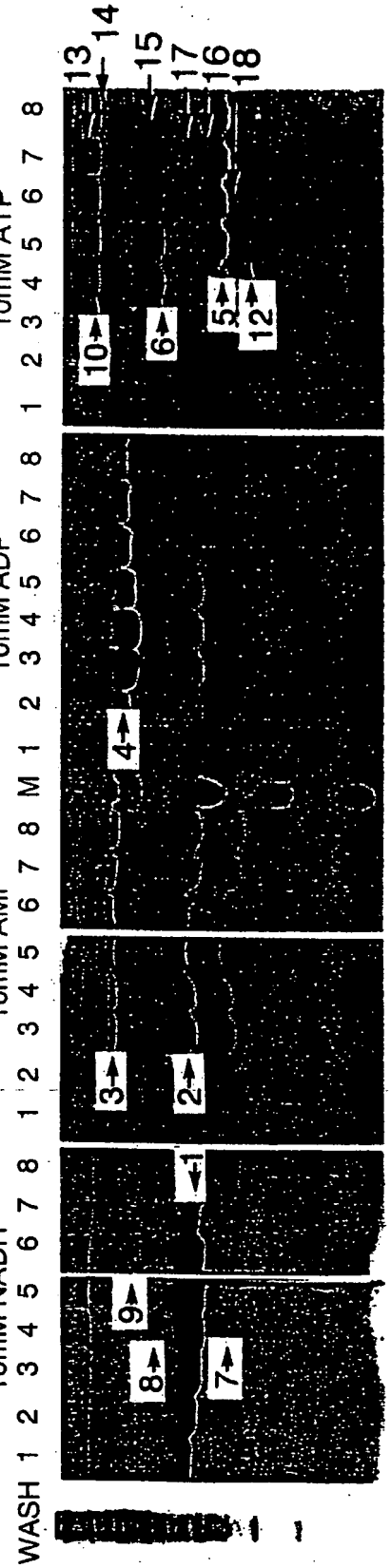

WASH 1 2
10mM NADH 3 4 5 6 7 8
10mM AMP 1 2 3 4 5 6 7 8
M 1 2
10mM ADP 3 4 5 6 7 8
10mM ATP 1 2 3 4 5 6 7 8
1
2
3
4
5
6
7
8
9
10
12
13
14
15
16
17
18

FIG. 5

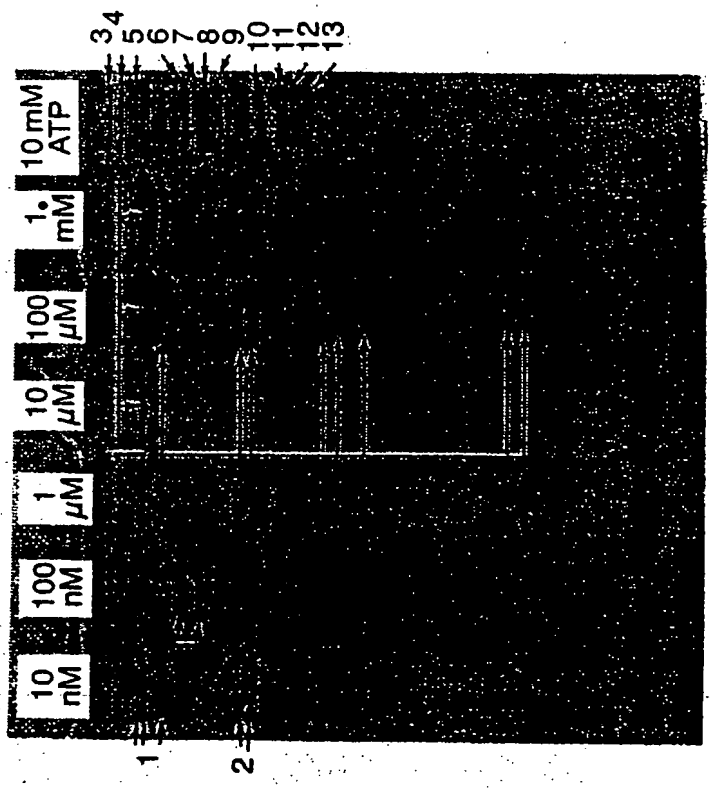


FIG. 6

SEQUENCE LISTING

<110> Haystead, Timothy A
University of Virginia Patent Foundation

<120> Proteome Mining

<130> 00164-02

<140>
<141>

<160> 4

<170> PatentIn Ver. 2.0

<210> 1
<211> 13
<212> PRT
<213> Oryctolagus cuniculus

<400> 1
Met Phe Phe Lys Asp Asp Ala Asn Asn Asp Pro Gln Trp
1 5 10

<210> 2
<211> 13
<212> PRT
<213> Oryctolagus cuniculus

<400> 2
Met Lys Ile Glu Phe Gly Val Asp Val Thr Thr Lys Glu
1 5 10

<210> 3
<211> 9
<212> PRT
<213> Oryctolagus cuniculus

<400> 3
Met Thr Lys Ala Asp Leu Ile Asn Asn
1 5

<210> 4
<211> 10
<212> PRT
<213> Oryctolagus cuniculus

<400> 4
Met Ile Gly Gln Phe Gly Val Gly Phe Tyr
1 5 10

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/09714

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) :G01N 33/53, 33/566, 33/543; C12Q 1/00; A61K 38/00

US CL : 435/4, 7.1, 7.92; 436/501, 518; 530/300, 344, 810, 811

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/4, 7.1, 7.92; 436/501, 518; 530/300, 344, 810, 811

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

NONE

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WEST 2.0, DERWENT, MEDLINE, SCISEARCH, BIOSIS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

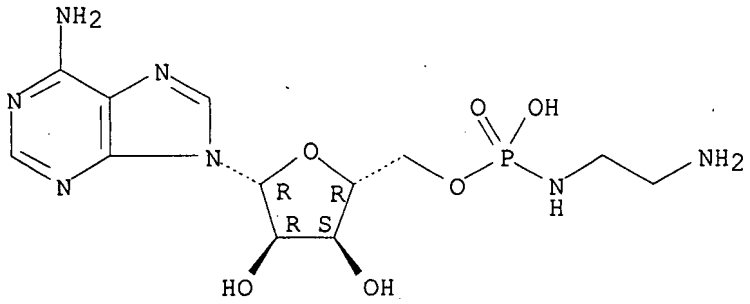
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y, P	BLUM et al. Isolation of Peptide Aptamers that Inhibit Intracellular Processes. PNAS. February 2000, Vol. 97, No. 5, pages 2241-2246, see entire document.	1-15
Y	HANDFILED et al. Strategies for Isolation of in Vivo Expressed Genes from Bacteria. FEMS Microbiology Reviews. January 1999, Vol. 23, No. 1, pages 69-91, see entire document.	1-15
Y	DAMER et al. Rapid Identification of Protein Phosphatase 1-binding Proteins by Mixed Peptide Sequencing Data Base Searching. The Journal of Biological Chemistry. 18 September 1998, Vol. 273, No. 38, pages 24396-24405, see entire document.	1-15

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

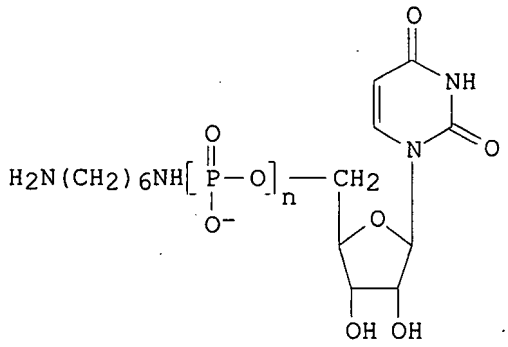
*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 11 JULY 2000	Date of mailing of the international search report 02 AUG 2000
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer P. PONNALURI Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet) (July 1998)★



L35 ANSWER 27 OF 34 CAPLUS COPYRIGHT 2006 ACS on STN
ACCESSION NUMBER: 1977:121683 CAPLUS
DOCUMENT NUMBER: 86:121683
TITLE: New affinity-chromatography adsorbents derived from uridine nucleotide phosphoryl amides
AUTHOR(S): Shibaev, V. N.; Kusov, Yu. Yu.; Kalinchuk, N. A.; Kochetkov, N. K.
CORPORATE SOURCE: N. D. Zelinskii Inst. Org. Chem., Moscow, USSR
SOURCE: Bioorganicheskaya Khimiya (1977), 3(1), 120-6
CODEN: BIKHD7; ISSN: 0132-3423
DOCUMENT TYPE: Journal
LANGUAGE: Russian
ED Entered STN: 12 May 1984
GI



I

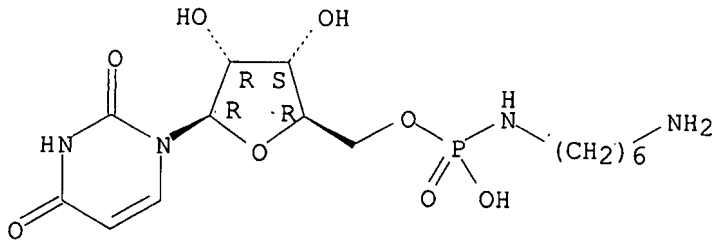
AB UDP and UMP condensed with Me3C6H2COCl to give mesitoic mixed anhydrides which were treated with H2N(CH2)6NH2 to give 63 and 53% I (n = 1,2). Treatment of the latter with BrCN-activated sepharose gave chromatog. adsorbents which contained immobilized UMP or UDP residues linked to the matrix through a phosphoamide bond.

IT **62149-09-1DP**, sepharose bound
RL: SPN (Synthetic preparation); PREP (Preparation)
(preparation and chromatog. adsorbent properties of)

RN 62149-09-1 CAPLUS

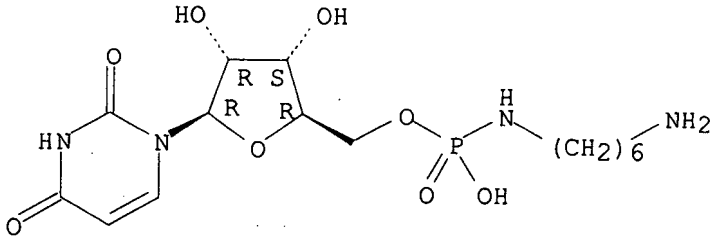
CN Uridine, 5'-[hydrogen (6-aminohexyl)phosphoramidate] (9CI) (CA INDEX NAME)

Absolute stereochemistry.



IT **62149-09-1P**
RL: RCT (Reactant); SPN (Synthetic preparation); PREP (Preparation); RACT (Reactant or reagent)
(preparation and reaction with cyanogen bromide-activated sepharose)
RN 62149-09-1 CAPLUS
CN Uridine, 5'-[hydrogen (6-aminohexyl)phosphoramidate] (9CI) (CA INDEX NAME)

Absolute stereochemistry.



L35 ANSWER 28 OF 34 CAPLUS COPYRIGHT 2006 ACS on STN
ACCESSION NUMBER: 1974:500920 CAPLUS
DOCUMENT NUMBER: 81:100920
TITLE: Possible role of crystals in the origins of life. VII. Adsorption and polymerization of phosphoramidates by montmorillonite clay
AUTHOR(S): Burton, F. G.; Lohrmann, R.; Orgel, L. E.
CORPORATE SOURCE: Salk Inst. Biol. Stud., San Diego, CA, USA
SOURCE: Journal of Molecular Evolution (1974), 3(2), 141-50
CODEN: JMEVAU; ISSN: 0022-2844
DOCUMENT TYPE: Journal
LANGUAGE: English
ED Entered STN: 12 May 1984
AB Nucleoside phosphoramidates derived from polyamines containing ≥3 amine groups are strongly adsorbed by Na and Mg montmorillonite clays even from very dilute solns. Heating the dried clay-phosphoramidate mixture results in the production of small amts. of the dinucleotides.
IT **52904-72-0**
RL: PEP (Physical, engineering or chemical process); PROC (Process)
(adsorption of, by montmorillonite)
RN 52904-72-0 CAPLUS
CN Adenosine, 5'-[hydrogen (2-aminoethyl)phosphoramidate] (9CI) (CA INDEX NAME)

Absolute stereochemistry.

ДЕМИЯ НАУК СССР

БИООРГАНИЧЕСКАЯ ХИМИЯ

М 3 * № 1 * 1977

БИООРГАНИЧЕСКАЯ ХИМИЯ
том 3 * № 1 * 1977

УДК 547.963.32 : 543.544

НОВЫЕ СОРБЕНТЫ ДЛЯ АФФИННОЙ ХРОМАТОГРАФИИ — ПРОИЗВОДНЫЕ ФОСФОАМИДОВ УРИДИНОВЫХ НУКЛЕОТИДОВ

Шибаев В. Н., Кусов Ю. Ю., Калинчук Н. А., Кочетков Н. К.

Институт органической химии им. Н. Д. Зелинского Академии наук СССР, Москва

Взаимодействием N^1-(уридин-5′-фосфорил)- и N^1-(уридин-5′-пирофосфорил)-1,6-диаминогексанов с активированной бромцианом сефарозой получены биоспецифические сорбенты для ферментов, взаимодействующих с уридинсодержащими нуклеотидами. Сорбенты содержат иммобилизованные остатки UMP и UDP, связанные с матрицей фосфоамидной связью.

Аффинная хроматография с использованием адсорбентов, содержащих иммобилизованные нуклеотиды, получила в последние годы широкое распространение для очистки разнообразных ферментов [1, 2]. Особенно большое внимание уделялось получению адсорбентов — производных адениновых нуклеотидов, в то время как получение носителей с иммобилизованными уридиновыми нуклеотидами, которые, по-видимому, могут найти широкое применение при очистке большой группы ферментов, взаимодействующих с уридинсодержащими нуклеотидами, изучено в значительно меньшей степени. В описанных в литературе примерах иммобилизация производных UMP на агарозных матрицах осуществлялась через остаток рибозы после периодатного окисления цис-гликольной группировки [3—5] или через остаток фосфорной кислоты с помощью *n*-аминофениловых эфиров [6, 7], а иммобилизация UDP — через остаток рибозы [4] или β-фосфата с помощью 6-аминогексилового эфира [8].

Недавно предложен новый метод иммобилизации нуклеотидов на носителях, основанный на взаимодействии смешанных ангидридов нуклеотидов и мезитиленкарбоновой кислоты с NH_2-группой аминогексилсефарозы [9, 10]. С помощью этого метода получены адсорбенты, связанные с различными моно- и олигонуклеотидами, а также нуклеотидполифосфатами, однако адсорбенты, содержащие производные уридиновых нуклеотидов, приготовлены не были.

Недостаток этого способа синтеза аффинных адсорбентов состоит в сохранении в полученной матрице незамещенных аминогрупп, что может приводить к появлению у сорбента ионообменных свойств и осложнениям при биоспецифической хроматографии [2]. Многообещающим кажется метод, основанный на взаимодействии лигандов, содержащих свободную аминогруппу, с активированной бромцианом сефарозой.

В настоящей работе мы сообщаем о получении на основе сефарозы адсорбентов, содержащих иммобилизованные остатки UMP и UDP, которые присоединены к матрице через остаток фосфата с помощью фос-

ХИМИЯ
ом 3 * № 1 * 1977

547.963.32 : 543.544

ТОГРАФИИ —
НУКЛЕОТИДОВ

ук Н. А.,

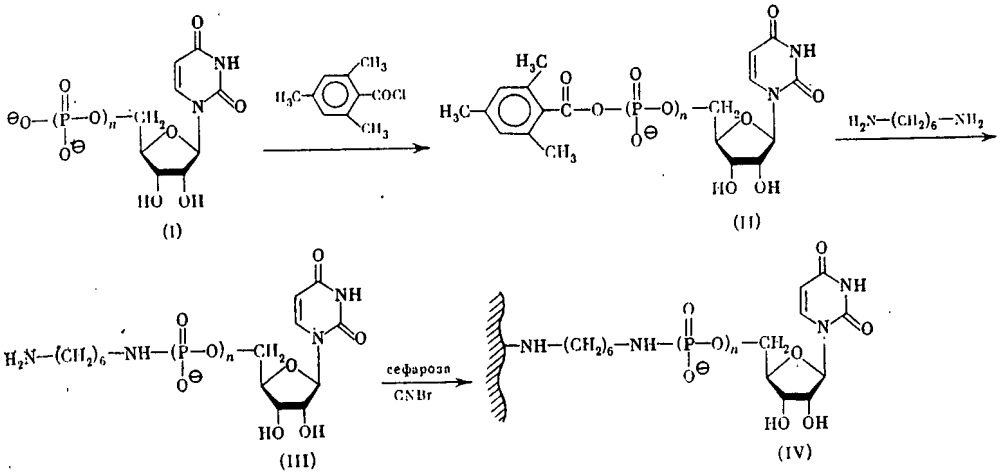
инского

ı-5′-пирофосфорил)-1,6-
лучены биоспецифичес-
одержащими нуклеоти-
и UDP, связанные с

бентов, содержащих
годы широкое рас-
в [1, 2]. Особенно
тов — производных
осителей с иммоби-
по-видимому, могут
группы ферментов,
ами, изучено в зна-
е примерах иммоби-
ах осуществлялась
ия цис-гликольной
ислоты с помощью
DP — через остаток
ового эфира [8].
уклеотидов на носи-
гидридов нуклеоти
аминогексилсефаро-
ы, связанные с раз-
отидполифосфатами,
новых нуклеотидов,

бентов состоит в со-
ногрупп, что может
свойств и осложне-
обещающим кажет-
содержащих свобод-
фарозой.
на основе сефарозы
ки UMP и UDP,
ата с помощью фос-

фоамидной связи [4]. Их получение из соответствующих нуклеотидов (I) включает в себя превращение последних в смешанные ангидриды (II) и затем в фосфоамиды — производные гексаметилендиамина (III), которые далее иммобилизуются на CNBr-сефарозе:



а: $n=1$; б: $n=2$

Конденсация UDP (Iб) с хлорангидридом мезитиленкарбоновой кислоты (MesCOCl) проводилась в условиях, использованных ранее для синтеза смешанных ангидридов [11, 12]. Анализ реакционной смеси с помощью хроматографии и электрофореза на бумаге показал присутствие двух главных нуклеотидных продуктов (IIа) и (IIб), образующихся в приблизительно равных количествах, неорганического фосфата и небольших количеств (Iа) и (Iб).

Этот результат указывает на значительное расщепление пирофосфатной связи нуклеотида в условиях реакции. В описанных ранее примерах получения смешанных ангидридов ADP и GDP такого расщепления пирофосфатной связи с образованием производных типа (IIа) не наблюдалось, хотя было отмечено появление неорганического фосфата (~15%) при взаимодействии MesCOCl с ATP [9].

Продукты реакции были разделены с помощью ионообменной хроматографии на DEAE-целлюлозе; смешанные ангидриды (IIа) и (IIб) были выделены с выходами 39 и 37% соответственно (см. таблицу).

Идентичность УФ-спектров исходных и полученных соединений свидетельствует об отсутствии изменений гетероциклического основания. Величины отношения нуклеозид: фосфор показывают, что (IIа) является производным UMP, а (IIб) — производным UDP. При кислотном гидролизе (IIа) и (IIб) образуется мезитиленкарбоновая кислота и соответствующие нуклеотиды. В спектре ПМР имеются сигналы, характерные для остатка мезитиленкарбоновой кислоты и уридина. Совокупность всех этих фактов позволяет однозначно приписать соединениям (IIа) и (IIб) структуры смешанных ангидридов мезитиленкарбоновой кислоты с уридин-5′-моно- и дифосфатом соответственно.

Взаимодействие (Iа) с хлорангидридом мезитиленкарбоновой кислоты гладко приводит к смешанному ангидриду (IIа), который после очистки ионообменной хроматографией получен с выходом 63%.

Для получения нуклеотидных лигандов со свободной аминогруппой смешанные ангидриды (IIа) и (IIб) вводили в реакцию с гексаметилендиамином, используя 20—30-кратный избыток последнего для предотвращения возможного образования симметричных бис-фосфоамидных производных. Реакция заканчивалась за 5—7 ч при 20° и за 20 ч при 5°. Фосфоамид (IIIа), продукт конденсации смешанного ангидрида (IIа) с гексаме-

Соединение	R_f в системах		R_{UMP} в буфере В	Отношение уридин : общий фосфор : кислотолабильный фосфор : аминогруппа
	А	Б		
(IIа)	0,80	0,78	0,60	1,00 : 1,03 : 0,00—
(IIб)	0,68	0,64	0,85	1,00 : 2,15 : 0,99—
(IIIа)	0,47	0,28	0,35	1,00 : 1,06 : 0,00 : 1,09
(IIIб)	0,29	0,18	0,65	1,00 : 2,14 : 1,07 : 1,07

тилендиамином, выделен из реакционной смеси ионообменной хроматографией на дауэксе 1 × 8: попытка использования более слабого анионита — DEAE-целлюлозы — была неуспешной. После обессоливания гель-фильтрацией на сефадексе G-15 выход (IIIа) составил 63%.

Структура соединения (IIIа) определена отношением нуклеозид : фосфор : аминогруппа (см. таблицу), а также данными кислотного гидролиза, в результате которого из фосфоамида (IIIа) образуется гексаметилендиамин и UMP.

Для выделения фосфоамида (IIIб) из реакционной смеси в качестве ионообменника использована DEAE-целлюлоза; соединение (IIIб) получено в виде триэтиламмониевой соли с выходом 53%. Аналитические данные (см. таблицу), а также спектр ПМР подтверждают его структуру.

Фосфоамиды (IIIа) и (IIIб) могут быть получены, исходя из UDP, без промежуточного выделения смешанных ангидридов (IIа) и (IIб). В таком варианте фосфоамид (IIIб) выделен ионообменной хроматографией на DEAE-целлюлозе с выходом 31%; по-видимому, фосфоамид (IIIа) может быть выделен из промывных вод ионообменной хроматографией на дауэксе 1 × 8. Полученный в одну стадию препарат (IIIб) содержал небольшое количество примеси, которая была отделена препаративным электрофорезом на бумаге.

Получение адсорбентов для аффинной хроматографии (IVа) и (IVб) осуществлено взаимодействием фосфоамидов (IIIа) и (IIIб) с активированной бромцианом сефарозой 4В. Для активации носителя применена модификация недавно описанной методики активации сефарозы в щелочном фосфатном буфере [13], состоящая в замене водного раствора бромциана раствором в ацетонитриле. Этот вариант активации геля оказался более эффективным, чем активация ацетонитрильным раствором бромциана в карбонатном буфере [14] *, и позволил получать воспроизводимые результаты.

Конденсация фосфоамидов (IIIа) и (IIIб) с активированной бромцианом сефарозой 4В проведена при pH 8,3. Остаточные активные центры сорбента удаляли обработкой этаноламином. В условиях реакции существенного гидролиза фосфоамидов не происходило, что позволяло их использовать для повторной конденсации. Были получены адсорбенты (IVа) и (IVб), содержащие 4,8 мкмоль фосфоамида (IIIа) и 5,2 мкмоль соединения (IIIб) на 1 мл геля соответственно.

Полученные адсорбенты оказались устойчивыми в условиях аффинной хроматографии при pH 8,1—8,5 и при хранении в нейтральных растворах при пониженной температуре; так, например, за 30—40 дней при 4° в растворе 1 М хлористого натрия степень десорбции УФ-поглощающего материала в случае адсорбента (IVа) составила менее 10%.

Степень иммобилизации полученного уридинфосфоамида сравнима с литературными данными по иммобилизации уридиновых нуклеотидов фосфоэфирной связью с помощью *n*-аминофениловых эфиров [6, 7] или в случае пирофосфатов с помощью 6-аминогексилового эфира [8].

* При активации геля в карбонатном буфере добавлением раствора бромциана в ацетонитриле, как это описано в работе [14], степень иммобилизации фосфоамидов не превышала 0,5—0,8 мкмоль/мл геля.

Таким образом, описанный в настоящей работе метод получения сорбентов с иммобилизованными уридиновыми нуклеотидами, связанными фосфоамидной связью с матрицей, позволяет просто и с высокой эффективностью осуществить присоединение остатков нуклеотидов к матрице, не содержащей свободных аминогрупп. Полученные адсорбенты могут оказаться полезными при выделении и исследовании ферментов, проявляющих сродство к уридиновым нуклеотидам. Наши предварительные опыты свидетельствуют о высокой эффективности сорбента (IVa) при очистке фермента, катализирующего перенос остатка галактозилфосфата с уридинфосфатгалактозы на полипренолфосфат при биосинтезе О-специфического полисахарида сальмонелл. Примененный в данной работе подход, очевидно, может быть распространен и на другие нуклеотиды.

Экспериментальная часть

В работе использовали UMP и UDP (Reanal, ВНР), три-*н*-октиламин (Schuchardt, ФРГ), сефадекс G-15 и сефарозу 4В (Pharmacia, Швеция), DEAE-целлюлозу (Whatman, Англия).

Аналитическую хроматографию и электрофорез проводили на бумаге FN-11 (Filtrak, ГДР), а препаративную — на Whatman 3 MM (Англия). При БХ использовали системы: А — этанол — 1 М ацетат аммония, 5 : 2, pH 7,5; Б — изопропанол — аммиак — вода, 7 : 1 : 2; электрофорез проводили в 0,05 М триэтиламмонийбикарбонатном буфере, pH 7,5 (буфер В) на приборе ПВЭФ-1 при градиенте напряжения 18 В/см.

Фосфорсодержащие и вещества с первичной аминогруппой обнаруживали на хроматограммах и электрофореграммах с помощью реагента Хайнса-Ишервуда и раствора нингидрина соответственно (прописи D45a и D112a в [15]). Количественное определение кислотолабильного и общего фосфора проводили по методам работ [16] и [17] соответственно, первичной аминогруппы — раствором тринитробензолсульфоната [18] с использованием для калибровки этаноламина.

Пиридин абсолютировали кипячением над P_2O_5, ацетонитрил — над CaH_2, бензол — над натрием и этанол — над магнием; все растворители перегоняли перед употреблением. Хлорангидрид мезитиленкарбоновой кислоты (MesCOCl) и бромциан получали по описанным методикам [19] и [20] соответственно.

Спектры ПМР снимали в D_2O на приборе WP-60 (Bruker, ФРГ) с накоплением в импульсном режиме (10 прохождений) с подавлением сигнала растворителя (диоксан в качестве внутреннего стандарта); УФ-спектры снимали на спектрофотометре Unicam SP-8000 (Англия).

Конденсация хлорангидрида мезитиленкарбоновой кислоты с UDP. Водный раствор уридин-5′-дифосфата (0,3 ммоль) пропускали через колонку (1 × 5 см) с дауэксом 50 (H^+-форма), нейтрализовали пиридином и упаривали до 1—2 мл. К остатку добавляли 0,45 ммоль три-*н*-октиламина в 1,5 мл пиридина, встряхивали до гомогенности, раствор упаривали и остаток тщательно высушивали многократной отгонкой со смесью абс. C_6H_6 — абс. EtOH (1 : 1). Полученный остаток растворяли в 3 мл смеси абс. C_5H_5N — абс. CH_3CN (3 : 1), добавляли 3,74 ммоль (0,6 мл) MesCOCl, энергично встряхивали в течение 8 мин при 20°, добавляли 12 мл воды, MesCOOH экстрагировали эфиром (6 × 5 мл); эфирный слой промывали водой (1 × 5 мл). Объединенный водный раствор разбавляли водой до 300 мл и наносили на колонку (3 × 24 см) с DEAE-целлюлозой (HCO_3^--форма); колонку промывали водой до отсутствия поглощения при 260 нм (390 мл) и элюировали линейным градиентом концентрации триэтиламмонийбикарбоната (ТЕАБ) от 0 до 0,4 М (по 1 л). Фракции, содержащие УФ-поглощающий материал, объединяли и лиофилизовали с добавлением воды (5 × 10 мл) для удаления ТЕАБ. Выход триэтиламмониевой соли (IIa) (элюируется 0,07 М ТЕАБ) составил 39% (117 мкмоль), а

триэтиламмониевой соли (IIб) (элюируется 0,14 М ТЕАБ) — 37% (111 мкмоль). Хроматографические и электрофоретические подвижности, а также аналитические данные приведены в таблице.

При гидролизе (0,1 н. HCl, 100°, 2 ч) ангидридов (IIа) и (IIб), по данным БХ и электрофореза, образуются MesCOOH, UMP, а также UDP в случае (IIб). Спектр ПМР аммониевой соли (IIб) (δ, м. д.): 7,69 д (6-H $J_{5,6}$ 8 Гц); 6,74 с (протоны бензольного ядра); 5,84 м (5-H + 1′-H); 2,15 с и 2,08 с (3 CH_3).

Смешанный ангидрид мезитиленкарбоновой кислоты с UMP (IIа). Три-*н*-октиламмониевую соль UMP (0,6 ммоль), полученную по вышеописанной методике, тщательно высушивали, растворяли в 10 мл абс. пиридина и обрабатывали 7,84 ммоль (1,2 мл) MesCOCl в условиях, аналогичных получению (IIб). Выход ангидрида (IIа) составил 63% (378 мкмоль). Аналитические данные для соединения (IIа), а также его хроматографические и электрофоретические характеристики приведены в таблице.

N^1-(уридин-5′-фосфорил)-1,6-диаминогексан (IIIа). К раствору 0,19 ммоль (IIа) в 2 мл воды при 20° прибавляли по каплям в течение 10 мин 6 ммоль свежеперегнанного гексаметилендиамина в 1,5 мл воды. Через 20 ч при 4° реакционную смесь разбавляли до 1 л, наносили на колонку (3 × 25 см) с DEAE-целлюлозой (HCO_3^--форма) и промывали водой (200 мл). Элюировали линейным градиентом раствора ТЕАБ от 0 до 0,5 М по 1 л. Основное вещество (~74%) оказалось в промывных водах, в то время как в первой фракции — всего около 12% соединения (IIIа). Во второй фракции обнаружено около 8% исходного UMP. Промывные воды и первую фракцию объединяли и наносили на колонку (1,5 × 18 см) с дауэксом 1 × 8 (HCO_3^--форма) при 4—10°. Колонку промывали водой (500 мл) и элюировали вещества с колонки растворами ТЕАБ: 0,05 (200 мл), 0,075 (200 мл), 0,1 (500 мл), 0,125 (250 мл) и 1 М (110 мл). Основное количество вещества (93%) смывалось с колонки 0,1 М ТЕАБ, эту фракцию многократно лиофилизовали с добавлением воды для удаления ТЕАБ. Остаток растворяли в 2,5 мл воды и обессоливали на колонке (1,5 × 75 см) с сефадексом G-15; фракцию, содержащую (IIIа) (63%), лиофилизовали. Аналитический образец (IIIа) (см. таблицу) получали препаративным электрофорезом на бумаге Whatman 3 MM в буфере В. При кислотном гидролизе соединения (IIIа) (0,1 н. HCl, 100°, 2 ч), по данным БХ и электрофореза, наблюдалось образование гексаметилендиамина и UMP.

N^1-(уридин-5′-пирофосфорил)-1,6-диаминогексан (IIIб). А. К раствору 55 мкмоль (IIб) в 0,5 мл воды добавляли по каплям в течение 10 мин раствор 1 ммоль свежеперегнанного гексаметилендиамина в 0,5 мл воды. Через 20 ч при 20° реакционную смесь разбавляли водой до 500 мл и наносили на колонку (3 × 23 см) с DEAE-целлюлозой (HCO_3^--форма); колонку промывали водой (300 мл) и линейным градиентом ТЕАБ pH 7,5, от 0 до 0,36 М (по 1 л). Фракции, содержащие УФ-поглощающие вещества, объединяли и многократно лиофилизовали с добавлением воды. Выход (IIIб) составил 29 мкмоль (53%). Аналитический образец (см. таблицу) получен препаративной БХ в системе А. Спектр ПМР аммониевой соли (IIIб) (δ, м. д.): 7,89 д (6-H, $J_{5,6}$ 8 Гц); 5,89 м (5-H + 1′H); 2,87 т (CH_2 при C_6 гексаметилендиамина); 1,245; 1,146; 1,02 и 0,92 (5 CH_2, при $C_{(1)}$—$C_{(5)}$ гексаметилендиамина).

Б. 0,6 ммоль три-*н*-октиламмониевой соли UDP растворяли в смеси 9 мл абс. пиридина и 1,5 мл абс. ацетонитрила, добавляли 1,2 мл (7,84 ммоль) MesCOCl и энергично встряхивали 8 мин при 20°. Добавляли 12 мл воды, экстрагировали эфиром (5 × 10 мл). В водный слой, упаренный до 1 мл, добавляли в течение 20 мин 6 ммоль свежеперегнанного гексаметилендиамина в 1,2 мл воды. Через 20 ч при 3° реакционную смесь разбавляли водой до 1 л и наносили на колонку (2,5 × 30 см) с DEAE-целлюлозой (HCO_3^--форма); колонку промывали водой (500 мл) и линейным градиентом ТЕАБ от 0 до 0,4 М по 1 л. Фракции, содержащие УФ-по-

лощающие вещества, объе
влением воды. Выход (III

Активация сефарозы 4B
мл 5 М фосфатного буфера.
о 4° и добавляли при перем
онитриле (0,1 мл). Через 1С
али на стеклянном фильтре
обавляли соответствующий

Иммобилизация фосфоам
ктивированной бромциано
обавляли 1 мл 0,25 М раств
IIIа), и оставляли при 4° н
ировали при 2000 об/мин,
иофилизовали и использо
еля добавляли 1 мл 0,25
ель переносили в колонк
),1 М CH_3COONa, 6 М моче
),5 М NaCl) и, наконец, 1
осфоамида (IIIа), определе
осфора, переходящего в р
ликвоты (0,1 мл) адсорбен
еля.

Иммобилизация фосфоа
вированной бромцианом се
амида (IIIб) по вышеописа
лиганда, определенное по
(10,3 мкмоль) в гидролиз
на 1 мл сефарозы.

Авторы глубоко благод
и А. С. Шашкову за съем

1. Спиридонова В. А. (1975)
г. 4 (под ред. Киселева Л
2. Lowe C. R., Dean D. G.
3. Robbertson D. L., Davids
4. Lamed R., Levin Y., Wil
5. Eichler D. C., Glitz D. G
6. Wilchek M., Corecki M.
7. Stewart G. R., Stevenson
8. Barker R., Olsen K. W.,
7135—7147.
9. Носова В. В., Вельмога
ев М. А. (1974) Докл. А
10. Носова В. В., Соколова
1133.
11. Соколова Н. И., Носова
АН СССР, 206, 129—132
12. Moon M. W., Khorana H
13. Sharma M., Slaunwhite
14. March S. C., Parikh J.,
15. Хайс И. М., Мацек К.
16. Кочетков Н. К., Будовс
АН СССР. Сер. хим., 89
17. Vaskovsky V. E., Kostet
129—141.
18. Satake K., Okuyama T.
19. Барнес Р. (1952) Си
Изд-во иностр. лит., М.
20. Хартман В., Дрегер Б.
Изд-во иностр. лит., М.

глощающие вещества, объединяли и многократно лиофилизовали с добавлением воды. Выход (IIIб) — 184 мкмоль (31%).

Активация сефарозы 4В бромцианом. К 1 мл сефарозы 4В добавляли 1 мл 5 М фосфатного буфера, pH 11,9, полученную суспензию охлаждали до 4° и добавляли при перемешивании раствор бромциана (100 мг) в ацетонитриле (0,1 мл). Через 10—15 мин перемешивания при 4° гель промывали на стеклянном фильтре 10 мл 0,25 М раствора $NaHCO_3$ и немедленно добавляли соответствующий лиганд.

Иммобилизация фосфоамида (IIIа) на сефарозе (IVа). К 1 мл свежеактивированной бромцианом сефарозы 4В в центрифужной пробирке добавляли 1 мл 0,25 М раствора $NaHCO_3$, pH 8,3, содержащего 103 мкмоль (IIIа), и оставляли при 4° на 20 ч при перемешивании. Смесь центрифугировали при 2000 об/мин, супернатант, содержащий 83 мкмоль (IIIа), лиофилизовали и использовали для повторной конденсации. К осадку геля добавляли 1 мл 0,25 М этаноламина и перемешивали 2 ч при 20°. Гель переносили в колонку и промывали последовательно 20 объемами 0,1 М CH_3COONa, 6 М мочевины, 0,1 М $NaHCO_3$ (все элюенты содержали 0,5 М NaCl) и, наконец, 1 М NaCl. Степень ковалентного присоединения фосфоамида (IIIа), определенная по УФ-поглощению и количеству общего фосфора, переходящего в раствор после гидролиза (0,6 н. HCl, 20°, 16 ч) аликвоты (0,1 мл) адсорбента (IVа), составила 4,8 мкмоль (IIIа) на 1 мл геля.

Иммобилизация фосфоамида (IIIб) на сефарозе (IVб). 1 мл свежеактивированной бромцианом сефарозы 4В обрабатывали 73,0 мкмоль фосфоамида (IIIб) по вышеописанной методике. Количество иммобилизованного лиганда, определенное по УФ-поглощению (5,2 мкмоль) и общему фосфору (10,3 мкмоль) в гидролизате адсорбента, составило 5,2 мкмоль (IIIб) на 1 мл сефарозы.

Авторы глубоко благодарны Н. И. Соколовой за ценные обсуждения и А. С. Шашкову за съемку спектров ПМР.

ЛИТЕРАТУРА

1. Спиридонова В. А. (1975) в сб. Итоги науки и техники. Молекулярная биология, т. 4 (под ред. Киселева Л. Л.), с. 176—217, ВИНИТИ, М.
2. Lowe C. R., Dean D. G. (1974) Affinity Chromatography, London.
3. Robbertson D. L., Davidson N. (1972) Biochemistry, 11, 533—537.
4. Lamed R., Levin Y., Wilchek M. (1973) Biochim. et biophys. acta, 304, 231—235.
5. Eichler D. C., Glitz D. G. (1974) Biochim. et biophys. acta, 335, 303—317.
6. Wilchek M., Gorecki M. (1969) Eur. J. Biochem., 11, 491—494.
7. Stewart G. R., Stevenson K. J. (1973) Biochem. J., 135, 427—441.
8. Barker R., Olsen K. W., Shaper J. H., Hill R. Z. (1972) J. Biol. Chem., 247, 7135—7147.
9. Носова В. В., Вельмога И. С., Соколова Н. И., Шабарова З. А., Прокофьев М. А. (1974) Докл. АН СССР, 219, 1134—1136.
10. Носова В. В., Соколова Н. И., Шабарова З. А. (1975) Биоорган. химия, 1, 1130—1133.
11. Соколова Н. И., Носова В. В., Шабарова З. А., Прокофьев М. А. (1972) Докл. АН СССР, 206, 129—132.
12. Moon M. W., Khorana H. G. (1966) J. Amer. Chem. Soc., 88, 1805—1809.
13. Sharma M., Slaunwhite W. R. (1975) Anal. Biochem., 68, 79—86.
14. March S. C., Parikh J., Cuatrecasas P. (1974) Anal. Biochem., 60, 149—152.
15. Хайс И. М., Мацек К. (1962) Хроматография на бумаге, Изд-во иностр. лит., М.
16. Кочетков Н. К., Будовский Э. И., Шибаев В. Н., Лебедева К. С. (1969) Изв. АН СССР, Сер. хим., 897—903.
17. Vaskovsky V. E., Kostetsky E. Y., Vasendin V. M. (1975) J. Chromatogr., 114, 129—141.
18. Satake K., Okuyama T., Ohashi M., Sinoda T. (1960) J. Biochem., 47, 654—660.
19. Барнес Р. (1952) Синтезы органических препаратов, сб. 3, с. 462—464, Изд-во иностр. лит., М.
20. Хартман В., Дрегер Б. (1949) Синтезы органических препаратов, сб. 2, с. 123, Изд-во иностр. лит., М.

Поступила в редакцию
1.VII.1976

NEW AFFINITY ADSORBENTS DERIVED FROM URIDINE NUCLEOTIDE PHOSPHOAMIDATES

SHIBAEV V. N., KUSOV Yu. Yu., KALINCHUK N. A., KOCHETKOV N. K.

N. D. Zelinsky Institute of Organic Chemistry, Academy of Sciences of the USSR, Moscow

Biospecific adsorbents for uridine nucleotide-linked enzymes were prepared through reaction of N^1-(uridine-5′-phosphoryl)-or N^1-(uridine-5′-pyrophosphoryl)-1,6-diaminohexanes with BrCN-activated sepharose. The adsorbents contained immobilized uridine-5′-phosphate or-pyrophosphate residues linked to matrix through phosphoamide bond

---

БИС

ФИЗИК
КОВАЛЕН

Иванова Г.

Всесоюзный научно и

Сопоставлен хар
на, присоединенног
винилпирролидона
фермента приводит
гурация модифицирс
шей скоростью, чем
фермента по сравне
область, что, по-ви
группы трипсина, у
ция трипсина прив
ингибиторам сывор

Иммобилизаци
ральных матрица
катализаторов, п
гических задач
связанные с вод
видимому, будут
сопротивлением,
ные производные
ственно в медици
в которых сочет
ванным ферментс
ностью введения

Цель настоящ
растворимого по
мическом связыв
тоновой кислоты
(*n*-толуолсульфо
диимид).

Химическое
в присутствии ка
между карбокси
та — по-видимо
α-аминогруппам
провождается ин